Journal of
Clinical Psychology

JANUARY 1956
Volume XII Number 1

EDITORIAL BOARD

FREDERICK C. THORNE, *Editor*
Editorial Office: 5 Pearl Street
Brandon, Vermont

JERRY W. CARTER, JR.
U. S. Public Health Service

HOWARD F. HUNT
University of Chicago

WILLIAM A. HUNT
Northwestern University

ELAINE F. KINDER
New York State Department of Mental Hygiene

C. M. LOUTTIT
Wayne University

ROBERT I. WATSON
Northwestern University

ELEANOR C. THORNE, *Business Manager*

Journal of Clinical Psychology
5 Pearl Street · · Brandon, Vermont

COMPANION VOLUMES

PRINCIPLES OF PSYCHOLOGICAL EXAMINING

by

FREDERICK C. THORNE, M.D., Ph.D.
Editor, Journal of Clinical Psychology

PRICE: $7.50

PRINCIPLES OF PERSONALITY COUNSELING

by

FREDERICK C. THORNE, M.D., Ph.D.
Editor, Journal of Clinical Psychology

PRICE: $6.00

SYSTEMATIC: The first comprehensive system of diagnosis and psychotherapy based solidly on basic psychological science.

THEORETICAL: Presenting a comprehensive theory of the nature of personality integration with special emphasis on the differential diagnosis and modification of the primary etiologic factors which actually organize socially significant behaviors.

ECLECTIC: All pertinent theories and facts in basic science psychology and clinical psychological science are related in a systematic applied integrative psychology.

COMPLETE: For all teaching and reference purposes.

ADVANCED: Both of these books present a new orientation to diagnosis, counseling and psychotherapy which is more than just a rehash of the literature. Principles of Psychological Examining is more than a textbook of psychometrics since tests and measures are presented only as they relate to more fundamental problems of psychological diagnosis. Similarly, Principles of Personality Counseling presents thorough discussions of the indications and contraindications for applying all known methods of psychotherapy after a primary diagnosis of factors requiring modification has been accomplished.

ORDER NOW

JOURNAL OF CLINICAL PSYCHOLOGY
5 Pearl Street Brandon, Vermont

JOURNAL OF CLINICAL PSYCHOLOGY

VOL. XII JANUARY 1956 No. 1

CONTENTS

Page

Multiple Classification by the Method of Least Squares. Paul Horst, University of Washington 1

Intelligence Test Results and Observations of Personality Disorder among 3594 Unwed Mothers in Minnesota. John S. Pearson and Phyllis L. Amacher, Minnesota Department of Public Welfare 16

Factors of Importance for Creativity. John E. Drevdahl, University of Minnesota Medical School 21

The Objective Rorschach. P. O. O'Reilly, Saskatchewan Hospital 27

How Important is Psychotherapy to the Hospitalized Psychiatric Patient? Bernard A. Stotsky, Veterans Administration Hospital, Brockton, Mass. 32

Selection of Abbreviated TAT Sets. Richard H. Dana, St. Louis State Hospital 36

The Relation between Certain Parent's Attitudes toward Child Rearing and the Direction of Aggression of their Young Adult Offspring. Frieda L. Bornston, Los Angeles City College, and James C. Coleman, University of California at Los Angeles 41

Two Types of Perceptual Experience. Noel Jenkin. University of New Brunswick 44

A Conversion Table for Relating the MMPI Group and Individual Form Items. Charles O. Gaston, University of Houston, Earl S. Taulbee, University of Nebraska College of Medicine, Maurice R. Seaquist, USAF, and Saul B. Sells, Randolph Field, Texas 49

Group Psychotherapy in Conjunction with a Physical Treatment. Richard C. Cowden, Melvin Zax and J. A. Sproles, Veterans Administration Hospital, Gulfport, Miss. 53

Score Equivalence of the Wechsler-Bellevue Intelligence Scales, Forms I and II. John R. Barry, Samuel C. Fulkerson, Albert L. Kubala and Maurice R. Seaquist, USAF School of Aviation Medicine, Randolph Field, Texas 57

Unrealistic Reporting as a Function of Extraverted Neurosis. J. D. Keehn, American University of Beirut 61

A Transposed Factor Analysis of a Group of Schizophrenic Patients. Wilson H. Guertin, Spring Grove State Hospital and Richard L. Jenkins, Veterans Administration, Washington, D. C. 64

Measures of Intelligence in Brain Damaged Children. James J. Gallagher, E. Paul Benoit and Herbert F. Boyd, University of Illinois 69

A Statistical Study of Freud's Theory of Personality Types. Ross Stagner, University of Illinois and J. Weldon Moffitt, Brigham Young University 72

Reliability of the Despert Fables. Helen E. Peixotto, Catholic University of America 75

The Interdependence of Ratings of Case Histories. William A. Deterline and A. W. Bendig, University of Pittsburgh 79

(Over)

CONTENTS—Continued

A Bender-Gestalt Diagnostic Validity Study. John A. Bowland, University of Florida and Herdis L. Deabler, VA Hospital, Gulfport, Miss. 82
Rorschach Card Rejection in Normal and Psychiatric Groups. Boyd D. Sisson, University of Nebraska, Earl S. Taulbee, University of Nebraska College of Medicine and Charles O. Gaston, University of Houston 85
A Comparison of the Ability of Normal and Brain Injured Subjects to Produce Additional Responses on a Second Administration of the Rorschach Test. Leonard Weiner and Bert Kaplan, University of Kansas, and Earl Brown, Wadsworth VA Hospital 89
Galvanic Skin Conductance as a Function of Successive Interviews. Barclay Martin, University of Wisconsin 91
Manifest Anxiety in Psychotics. Harold Rubin, Leo Schneiderman, William C. Hallow and Robert J. Jones, Veterans Administration Hospital, Lebanon, Penna. 94

Editorial Opinion 97
Book Notices 99

GENERAL INFORMATION

The *Journal of Clinical Psychology* is an independent journal dedicated to the advancement of the clinical method in psychology. Although primarily a scientifically oriented professional journal limited to publication of original research reports and authoritative theoretical articles, it aims to foster the promotion and expansion of clinical psychology as an applied science.

Communications and articles submitted for publication are welcome and should be addressed to the Editorial Office, 5 Pearl Street, Brandon, Vermont.

ORIGINAL ARTICLES are published only with the understanding that they are contributed exclusively to the *Journal of Clinical Psychology*. The publishers are not responsible for statements made or opinions expressed by contributors in articles printed in these columns. Copyrights cover publication in the *Journal of Clinical Psychology* and articles may not be reproduced without permission of the publishers.

THE STYLE of all manuscripts including bibliographies must be in accordance with the instructions given in: Anderson, J. E., and Valentine, W. L. The preparation of articles for publication in the journals of the American Psychological Association. *Psychol. Bull.*, 1944, 41, 345-376. References should be arranged alphabetically according to the directions given above. Full address of the author should appear somewhere in the article, preferably at the end.

REPRINTS are furnished on order only and must be requested on the order blank provided when galley proofs are returned. Change of address notices should be sent one month before moving and should include old and new addresses. Undelivered copies will be remailed only at subscribers' expense.

The *Journal of Clinical Psychology* is published quarterly, appearing in January, April, July and October, at Rutland, Vermont, by Frederick C. Thorne, M.D. Subscription price $7.50; Canadian subscriptions $8.00; foreign subscriptions $8.50 including postage; single copies $2.00. Entered as second class matter January 15, 1945, at the post office at Burlington, Vermont, under the Act of March 3, 1879. Reentered at the post office at Brandon, Vermont. Additional entry at the post-office at Rutland, Vermont. *Title registered, U. S. Patent Office.* Editorial and Business Offices, 5 Pearl Street, Brandon, Vermont.

Copyright, 1956, Frederick C. Thorne, M. D.

MULTIPLE CLASSIFICATION BY THE METHOD OF LEAST SQUARES[1]

PAUL HORST

University of Washington

INTRODUCTION

The problem of classifying individuals into one of two or more specified groups on the basis of a set of observed measures has been of increasing interest to psychologists in recent years. Clinical psychologists, for example, have been concerned with techniques for utilizing a set of measurements on an individual as an aid in deciding into which of a specified set of groups he should be classified, such as, normal, anxiety state, hysteria, etc.

Essentially the approach to this problem begins with a group of persons, each of which has already been classified into one of a number of subgroups on the basis of expert psychiatric or psychological judgment. For each person it is assumed that a score is available on each of a number of psychological tests. The problem is then to find a set of weights corresponding to each subgroup, by means of which a linear combination of the test scores is obtained for each person. The weights corresponding to each subgroup are to be determined so that the weighted score will "best" discriminate the persons in the corresponding subgroup from persons in other subgroups. The presumption is that these weights established on persons whose classification is already known may subsequently be used as a basis for classifying individuals whose classification has not yet been determined.

A great deal of rather elaborate statistical theory has been developed as a basis for determining these "linear discriminant functions". However, it is not the purpose of this article to review the literature on this subject. Access to this literature can be obtained for example through the references given by Rao[4] and Anderson[1] including the pioneer work of R. A. Fisher and Neyman and Pearson. The method for determining the linear discriminant weights outlined by Rao[4] is probably typical of currently approved methods. In this article we shall briefly outline the method recommended by him, and then propose a method which is based on what we regard as a model more simple and conventional than that of Rao *et al.*

We use the same example as Rao[4] taken from Rao and Slater[3]. Scores on three tests, A, B, and C, of 256 army recruits classified by their neurotic condition have the mean values as given in Table 1. Rao's within-group variance-covariance matrix and its inverse are given in Table 2.

TABLE 1. MEAN SCORES OF NEUROTIC GROUPS (RAO AND SLATER, 1949)

Group	Sample Size	Mean Score A	B	C
Anxiety State	114	2.9298	1.1667	0.7281
Hysteria	33	3.0303	1.2424	0.5455
Psychopathy	32	3.8125	1.8438	0.8125
Obsession	17	4.7059	1.5882	1.1176
Personality Change	5	1.4000	0.2000	0.0000
Normal	55	0.6000	0.1455	0.2182
Total	256			
Test Means, Total Sample		2.6406	1.0508	0.6172

[1]This research was carried out under Research Grant M-743 MH (1) from the Public Health Service, National Institutes of Health. Much credit is due Charlotte MacEwan for major editorial and computational assistance. Elizabeth Cross deserves a great deal of credit for her special competence in typing the manuscript. Computational assistance and valuable suggestions were also provided by Robert E. Dear, Ivar Lovaas, William M. Meredith, and Donald F. Mills. Over-all supervision of the research and preparation of the report was provided by William V. Clemans.

TABLE 2. WITHIN-GROUP VARIANCE-COVARIANCE MATRIX (λ_{ij}), AND ITS INVERSE (λ^{ij})

	(λ_{ij})			(λ^{ij})		
	A	B	C	A	B	C
A	2.300851	0.251578	0.474169	0.543234	−0.200195	−0.420813
B	0.251578	0.607466	0.035774	−0.200195	1.725807	0.055767
C	0.474169	0.035774	0.595094	−0.420813	0.055767	2.012357

For a particular group Rao gives the linear discriminant score as:

$$l_1A + l_2B + l_3C - \tfrac{1}{2}(l_1m_1 + l_2m_2 + l_3m_3) \quad ,$$

where

$$l_1 = \lambda^{11}m_1 + \lambda^{12}m_2 + \lambda^{13}m_3 \quad ,$$
$$l_2 = \lambda^{21}m_1 + \lambda^{22}m_2 + \lambda^{23}m_3 \quad ,$$
$$l_3 = \lambda^{31}m_1 + \lambda^{32}m_2 + \lambda^{33}m_3 \quad ,$$

m_1, m_2, and m_3 are the mean values of A, B, C, and the elements λ^{ij} come from the inverse of the within group variance-covariance matrix. For example in the case of the anxiety state group, for the test means, we have from Table 1:

$$m_1 = 2.9298 \qquad m_2 = 1.1667 \qquad m_3 = 0.7281 \quad .$$

For the l's or discriminant weights we have

$$l_1 = 0.5432(2.9298) - 0.2002(1.1667) - 0.4208(0.7281) = 1.0515,$$
$$l_2 = -0.2002(2.9298) + 1.7258(1.1667) + 0.0558(0.7281) = 1.4676,$$
$$l_3 = -0.4208(2.9298) + 0.0558(1.1667) + 2.0124(0.7281) = 0.2975.$$

The constant term is given by:

$$\tfrac{1}{2}(l_1m_1 + l_2m_2 + l_3m_3) = \tfrac{1}{2}[1.0515(2.9298) +$$
$$+ 1.4676(1.1667) + 0.2975(0.7281)] = 2.5047.$$

Finally the discriminant score for the anxiety state is, according to Rao

$$L = 1.0515A + 1.4676B + 0.2975C - 2.5047.$$

Rao points out that for purposes of classification the expression to be calculated is $L + \log_e\pi$ where π is the relative frequency of a particular group. Since the population proportions are not known, Rao uses the proportions from the experimental data. The linear discriminant weights together with the constant terms given by Rao are listed in Table 3.

TABLE 3. THE LINEAR DISCRIMINANT SCORES FOR VARIOUS GROUPS

Groups	Coefficients of Measurements			Constant Term	
	A	B	C	(a) In Terms of General Proportion	(b) For proportions from Present Data*
Normal	0.2050	0.1431	0.1947	$-0.0931 + \log_e\pi_1$	−1.6311
Personality Change	0.7204	0.0649	−0.5780	$-0.5107 + \log_e\pi_2$	−4.4465
Anxiety State	1.0515	1.4676	0.2974	$-2.5047 + \log_e\pi_3$	−3.3137
Hysteria	1.1678	1.5679	−0.1081	$-2.7139 + \log_e\pi_4$	−4.7626
Psychopathy	1.3599	2.4641	0.1336	$-4.9182 + \log_e\pi_5$	−6.9977
Obsession	1.7680	1.8611	0.3573	$-5.8375 + \log_e\pi_6$	−8.5495

*$\pi_1 = 0.21484$, $\pi_2 = 0.01953$, $\pi_3 = 0.44531$, $\pi_4 = 0.12891$, $\pi_5 = 0.12500$, $\pi_6 = 0.06641$.

Rao[4] states that the rule of procedure is to assign an individual to the group for which $L + \log_e \pi$ is highest and that if the population proportions are not known, the maximum likelihood method of estimation leads to the rule of assigning an individual to the group for which L is highest.

It can be shown (Appendix D) that if we place a single individual's classification scores as calculated by Rao's method in rank order, this rank order will be independent of the origins for the test scores. Therefore, if a person is classified in the group for which his classification score is highest, the same results would be obtained whether one used deviation or raw score means by Rao's method. But if we consider the classification scores corresponding to a given classification for the total sample the rank order of these scores is not in general the same when raw score means rather than deviation score means are used in Rao's method.

Obviously, if one is concerned with comparing two persons' classification scores obtained by Rao's method for a given category, the question of using deviation or raw score means is not a matter of indifference. Rao[4] does not explicitly consider the question of origin of measurement. However, it is evident from Table 1 that the test score means are not obtained from deviation scores.

The method of group discrimination we shall propose is based on a conventional least square model. As is well known, regression weights in the traditional sense are independent of the origins of measurement of the independent variables even though the constant terms are not. Since the origin of measurement for test scores is in general arbitrary, any method to be acceptable should yield weights which are independent of these origins. If Rao's weights were based on means from deviation scores they would be independent of any arbitrary origins. Therefore, we have recalculated Rao's weights using deviation score means, so that corresponding to Rao's weights in Table 3 we now have Table 4. This provides a more adequate basis of comparison with the method to be outlined.

TABLE 4. LINEAR DISCRIMINANT SCORES FOR VARIOUS GROUPS WHEN DEVIATION SCORE MEANS ARE USED

Group	A	B	C
Anxiety State	.0893	.1519	.1106
Hysteria	.2084	.2546	−.3047
Psychopathy	.4052	1.1723	−.0571
Obsession	.8231	.5549	.1720
Personality Change	−.2498	−1.2845	−.7860
Normal	−.7777	−1.2043	.0052

Although it is not our purpose to review the traditional statistical rationale on which the method outlined by Rao[4] is based, nevertheless, for comparing the approach considered here with Rao's method, it may be instructive to outline a rationale which is mathematically somewhat simpler and which yields discriminant weights proportional to those obtained by Rao's method when deviation test score means are used.

We define an optimal set of discriminant weights for a given group as that set which yields a set of classification scores satisfying the following condition: The squared distance of the mean classification score for the particular group from the mean of the same scores for all of the groups is a maximum.

We must, however, define a unit of measurement in terms of which this squared distance is measured. Two rather obvious ways in which this unit of measurement may be defined suggest themselves. First we may define the unit of measurement corresponding to a particular set of classification scores as the weighted average variance of the groups for these scores. The weight for each group is taken as the ratio of the sub-sample size to the total sample size. The sum of the weights would then be unity. A second rather obvious unit of measurement is the variance of a particular set of classification scores for all the groups combined. If we adopt the first definition

of a unit of measurement then the optimal discriminant weights for a given group are proportional to the weights obtained by Rao's method when deviation test scores are used (Appendix A). If we adopt the second definition of a unit of measurement we get a set of discriminant weights for each group which is a special case of conventional least square regression weights (Appendix B). This is the procedure alternate to Rao's method which we shall describe.

The Method

We begin by regarding the test scores as a set of independent variables. We have as many dependent variables as we have groups. A person's score for the dependent variable corresponding to a particular group is unity if he has been assigned to a group and zero if not. Therefore, the problem is to determine for each group a set of regression weights, which, when applied to the matrix of test scores, will reproduce as closely as possible, in the least square sense, the "0" and "1" values in the set of "scores" for the corresponding dependent or criterion variable. It can readily be shown (Appendix C), that if there are only two groups then the conventional multiple regression method and the method of Rao *et al.* yield weights which are proportional, although the constant term will not in general be the same. However, in the case of more than two groups the weights obtained by the two methods will not in general be proportional.

The steps in the computational procedure are as follows (Appendix F):

1. We begin, as does Rao, with a matrix of raw score means for each group on each variable as in Table 1.

2. Next we subtract the total group mean for a given test from each of the group means. This gives the matrix of deviation score group means shown in Table 5. For example, the first entry in the first column of Table 5 is 2.9298 - 2.6406 = .2892.

Table 5. Deviation Score Group Means

Group	A	B	C
Anxiety State	.2892	.1159	.1109
Hysteria	.3897	.1916	−.0717
Psychopathy	1.1719	.7930	.1953
Obsession	2.0653	.5374	.5004
Personality Change	−1.2406	−.8508	−.6172
Normal	−2.0406	−.9053	−.3990

3. The deviation score group means for each test are multiplied by the corresponding P value, i.e., the number in the group divided by the total number. These values are respectively .4453, .1289, .1250, .0664, .0195, .2148. The results are given in Table 6. For example, the first entry in column 1 of Table 6 is .2892 x .4453 = .1288.

Table 6. Deviation Score Group Means Multiplied by the Corresponding Proportion of Individuals in the Group

Group	A	B	C
Anxiety State	.1288	.0516	.0494
Hysteria	.0502	.0247	−.0092
Psychopathy	.1465	.0991	.0244
Obsession	.1371	.0357	.0332
Personality Change	−.0242	−.0166	−.0120
Normal	−.4383	−.1945	−.0857

4. The matrix of test score variances and covariances for the total group is computed. This is given in Table 7.

TABLE 7. TEST SCORE VARIANCES AND COVARIANCES

	A	B	C	Σ
A	3.6833	.8776	.7609	5.3218
B	.8776	.8920	.1640	1.9336
C	.7609	.1640	.6503	1.5752

5. The next step is to premultiply the transpose of the matrix in Table 6 by the inverse of the matrix in Table 7. It is, however, not necessary first to find the inverse of the matrix in Table 7. Methods for performing such multiplications without first finding the inverse are well known[2]. The computations in this step give the sets of weights to be applied to the raw scores. These are shown in Table 8.

TABLE 8. LINEAR DISCRIMINANT WEIGHTS AND CONSTANT TERMS FOR METHOD OF LEAST SQUARES

Group	A	B	C	K
Anxiety State	.0176	.0318	.0473	.3362
Hysteria	.0175	.0177	−.0391	.0882
Psychopathy	.0192	.0938	−.0086	−.0190
Obsession	.0340	.0047	.0101	−.0346
Personality change	.0003	−.0162	−.0148	.0449
Normal	−.0887	−.1318	.0052	.5843

6. It is necessary also to calculate a constant term for each group. This is done by subtracting from the P value for a given group the product sum of total group means by the corresponding weights, from Table 8, for the given group. For example the constant term for the anxiety state is

$$.4453 - (2.6406 \times .0176 + 1.0508 \times .0318 + .6172 \times .0473) = .3362.$$

The constant terms are given in the last column of Table 8.

As a basis of comparison we shall calculate for each method (i.e., for the method of least squares, and for Rao's method when deviation score means are used), the correlation of each set of classification scores with its corresponding criterion, i.e., with the dichotomous score of assignment or non-assignment to the corresponding group. These correlations are, in fact, point biserial correlations between estimated classification scores and actual two-valued classification scores. It is not, however, necessary actually to calculate the estimated classification scores for the persons in the experimental groups in order to get these correlations (Appendix E). The correlations are calculated as follows:

7. Product sums for each row in Table 7 with each row of Table 4 are calculated to give Table 9. For example the first entry is

$$3.6833 \times .0893 + .8776 \times .1519 + .7609 \times .1106 = .5464.$$

8. Product sums for each column in Table 9 with the corresponding row of Table 4 are calculated to give the first row of Table 10. For example the first entry is

$$.5464 \times .0893 + .2320 \times .1519 + .1648 \times .1106 = .1023.$$

TABLE 9. MATRIX $G_T h_W$: MATRIX IN TABLE 7 POSTMULTIPLIED BY THE TRANSPOSE OF THE MATRIX IN TABLE 4.

	Anxiety State	Hysteria	Psychopathy	Obsession	Personality change	Normal
A	.5464	.7592	2.4778	3.6496	−2.6454	−3.9174
B	.2320	.3600	1.3919	1.2455	−1.4939	−1.7559
C	.1648	.0022	.4634	.8292	−.9119	−.7859

TABLE 10. COMPUTATIONS TO OBTAIN THE CORRELATIONS OF CLASSIFICATION SCORES WITH ACTUAL CLASSIFICATIONS

Computational Step	Group					
	Anxiety State	Hysteria	Psychopathy	Obsession	Personality change	Normal
1	.1023	.2492	2.6093	3.8377	3.2965	5.1571
2	.2470	.1123	.1094	.0620	.0191	.1687
3	.0253	.0280	.2855	.2379	.0630	.8700
4	.1591	.1673	.5343	.4878	.2510	.9327
5	.0248	.0196	.1741	.1384	.0368	.5747
6	.006244	.001675	.011899	.005165	.000439	.064067
7	.0253	.0149	.1088	.0833	.0230	.3798

9. Each P_i is multiplied by $1 - P_i$ and entered in row 2 of Table 10. For example the first entry is $.4453 \times .5547 = .2470$.

10. Each entry in row 1 of Table 10 is multiplied by the corresponding entry in row 2 to give the entries in row 3. For example the first entry is $.1023 \times .2470 = .0253$.

11. The square root of each entry in row 3 of Table 10 is entered in row 4. For example the first entry is $\sqrt{.0253} = .1591$.

12. Row 5 of Table 10 consists of the product sums of corresponding rows in Tables 4 and 6. For example the first entry is

$$.0893 \times .1288 + .1519 \times .0516 + .1106 \times .0494 = .0248.$$

13. Row 6 of Table 10 contains the product sums of corresponding rows from Tables 6 and 8. For example the first entry is

$$.1288 \times .0176 + .0516 \times .0318 + .0494 \times .0473 = .006244.$$

14. Each entry in row 6 of Table 10 is divided by the corresponding entry of row 2 of Table 10 to give the entries in row 7. For example the first entry is $(.006244) / (.2470) = .0253$.

Table 11 gives the point biserial correlations of classification scores with actual classification. These are obtained from Table 10 as follows:

15. Each entry in row 5 of Table 10 is divided by the corresponding entry in row 4 of Table 10 to get row 1 of Table 11. For example the first entry is $(.0248) / (.1591) = .1559$. This row gives the correlation of classification scores with actual classifications for Rao's method.

16. The square root of each entry in row 7 of Table 10 is entered in row 2 of Table 11. For example the first entry is $\sqrt{.0253} = .1591$. This row gives the correlations of the classification scores with the actual classification for the multiple regression method.

As may be noted from Table 11, for all groups except the "Normal" group, the multiple regression method yields coefficients slightly larger than those found in this example by Rao's method.

TABLE 11. CORRELATIONS OF CLASSIFICATION SCORES WITH ACTUAL CLASSIFICATION

	Group					
	Anxiety State	Hysteria	Psychopathy	Obsession	Personality change	Normal
Rao's Method	.1559	.1172	.3258	.2837	.1466	.6162
Multiple Regression Method	.1591	.1221	.3298	.2886	.1517	.6163

APPENDICES

A. Scaling with Weighted Average Variances

X is an N x n matrix of raw score measures where N is the number of cases and n the number of variables.

x is the same matrix for deviation scores.

L is an N x s classification matrix with a "1" in each row and all remaining elements "0". Each row of L corresponds to a person and each column to a classification category. A person has a score of "1" for the category in which he has been classified and a score of zero for all other categories. There are s categories.

P_i is the proportion of cases in group i.

D_p is a diagonal matrix of the P_i's.

M is an n x s matrix of raw score means for each group on each variable.

m is an n x s matrix of deviation score means on each variable.

h is an n x s matrix of discriminant weights.

Consider now

$$xh_i = y_i \quad , \tag{1}$$

where h_i is the i'th column vector of h to be determined. Let V_i be a vector of subgroup means for y_i. It can be shown that

$$V_i = D_p^{-1} \frac{L'y_i}{N} \quad . \tag{2}$$

Since x is in deviation scores we write the variance of y_i as

$$\sigma^2_{T_i} = \frac{y_i' \, {}^1 y_i}{N} \quad . \tag{3}$$

Let

$$\sigma^2_{B_i} = V_i' D_p V_i \quad . \tag{4}$$

Let

σ_{ij}^2 be the variance of group j on y_i .

In scalar notation we let

$$\sigma^2_{w_i} = \sum_j P_j \, \sigma_{ij}^2, \quad , \tag{5}$$

so that $\sigma^2_{w_i}$ is the weighted average of the group variances for y_i. For simplicity in algebraic manipulation we shall assume sufficiently large numbers in each subgroup so that the conventional corrections for degrees of freedom may be ignored.

It can be shown that

$$\sigma^2_{w_i} = \sigma^2_{T_i} - \sigma^2_{B_i} \quad . \tag{6}$$

Equation 6 gives the conventional relationships involving within, total, and between group variances.

Consider now the problem of maximizing the absolute distance between the mean of group i on y_i from the total mean for y_i. This latter is zero.

First we choose as our unit of measurement the weighted average group variance for y_i which is given by (5). This means that we maximize the ratio,

$$\phi_w = \frac{V_{ii}^2}{\sigma^2_{w_i}} \quad . \tag{7}$$

It can be proved that

$$m' = D_p^{-1} \frac{L'x}{N} \quad . \tag{8}$$

From (1) and (2) and (8)

$$V_B = m_{.i}' h_i \quad , \tag{9}$$

where $m_{.i}'$ is the ith column of m expressed as a row vector.

Let

$$\frac{x'x}{N} = G_T \quad , \tag{10}$$

$$mD_p m' = G_B \quad , \tag{11}$$

$$G_w = G_T - G_B \quad . \tag{12}$$

From (1), (2), (3), (4), (6), (8), (10), (11), and (12)

$$\sigma_w^2 = h_i' G_w h_i \quad . \tag{13}$$

Substituting (9) and (13) in (7),

$$\phi_w = \frac{h_i' (m_{.i} m_{.i}') h_i}{h_i' G_w h_i} \quad . \tag{14}$$

To maximize (14) we differentiate ϕ with respect to h_i' and equate to zero. This gives

$$(m_{.i} m_{.i}') h_i - \phi G_w h_i = 0, \tag{15}$$

or

$$h_i = K_w G_w^{-1} m_{.i} \quad , \tag{16}$$

where K_w is a constant. Since G_w is Rao's within group variance-covariance matrix, Equation 16 differs from Rao's formula only by the constant and the use of deviation score means, and shows that his discriminant weights are proportional to those obtained by maximizing the squared distance of the group mean from the total mean, using the weighted average group variance as a unit of measurement.

B. Scaling with Total Group Standard Deviation

If now we use the total group variance as a unit of measurement we have instead of Equation 14

$$\phi_T = \frac{h_i' (m_{.i} m_{.i}') h_i}{h_i' G_T h_i} \quad . \tag{17}$$

Using the preceding method we get

$$h_i = G_T^{-1} m_{.i} K_T \quad , \tag{18}$$

where K_T is a scalar quantity.

Since G_T is the total variance-covariance or dispersion matrix, h_i given by Equation 18 can be shown to be proportional to the method we have outlined. To show this we next derive the weights by conventional least square methods. We let X and L be the matrices of independent and dependent measures respectively. To get L in deviation form we let

$$(I - \frac{1\,1'}{1'\,1}) L = t \quad , \tag{19}$$

and note also that

$$(I - \frac{1\,1'}{1'\,1})X = x \quad , \tag{20}$$

where 1 is an Nth order column vector of "1"'s. We write

$$xh - t = \epsilon \quad . \tag{21}$$

The least square solution is known to be

$$h = (x'x)^{-1}\, x't \quad , \tag{22}$$

but it can be shown that

$$x't = NmD_p \quad . \tag{23}$$

From (10), (22) and (23)

$$h = G_T^{-1}\, mD_p \quad . \tag{24}$$

Equation 24 shows that the i'th column of h is proportional to h_i in Equation 18.

To solve for the constant terms we substitute (19) and (20) in (21)

$$(I - \frac{1\,1'}{1'\,1})Xh - (I - \frac{1\,1'}{1'\,1})L = \epsilon \quad . \tag{25}$$

Now

$$\frac{1'X}{1'\,1} = V_M' \quad , \tag{26}$$

where V_M' is a vector of total group raw score means. Also

$$\frac{1'L}{1'\,1} = V_p' \quad , \tag{27}$$

where V_p' is a row vector of the p_i values. Using Equations 26 and 27 in 25 gives

$$Xh + 1(V_p' - V_M'h) = L + \epsilon \quad . \tag{28}$$

The vector in parentheses in Equation 28 gives the constant terms corresponding to the various categories.

C. Proof that Rao's Method and the Method of Least Squares Yield Proportional Weights for Only Two Groups

If there are only two groups, it can readily be shown that

$$m = m_{.1}(1, -\frac{p_1}{q_1}) \quad . \tag{29}$$

Therefore

$$mDm' = m_{.1}\begin{bmatrix} 1, -p_1 / q_1 \end{bmatrix}\begin{bmatrix} p_1 & 0 \\ 0 & q_1 \end{bmatrix}\begin{bmatrix} 1 \\ -p_1/q_1 \end{bmatrix} m_{.1}' \quad ,$$

or

$$mDm' = \frac{p_1}{q_1}\, m_{.1}\, m_{.1}' \quad . \tag{30}$$

Using (30) in (12) we get

$$G_w = \left[G_T - \frac{p_1}{q_1}\, m_{.1}\, m_{.1}' \right] \quad . \tag{31}$$

Substituting (31) in (15) and taking $i = 1$ gives

$$m_{.1}m_{.1}'\,h_1 - \phi\left[C_T - \frac{p_1}{q_1}\,m_{.1}\,m_{.1}'\right]h_1 = 0 \quad ,$$

or

$$h_1 = G_T^{-1}\,m_{.1}C \quad , \tag{32}$$

where C is a scalar constant given by

$$C = \left[\frac{p_1}{q_1} + \frac{1}{\phi}\right]m_{.1}'\,h_1 \quad . \tag{33}$$

But the right side of Equation 18 for $i = 1$, and Equation 32 are the same except for a scalar constant which shows that the weights in the two equations are proportional. It can be readily shown that in the case of only two groups h_2 is proportional to $-h_1$ for either method. Therefore h_2 for Rao's method and the least square method must also be proportional.

D. The Effect of Change of Origin on the Classification Scores by Rao's Method

First we consider the equation for getting the classification scores from raw scores by Rao's method (4) when the weights are based on raw score means. In matrix notation this is

$$Xh - \tfrac{1}{2}1\,1'\left[D_{M'h} - 2D_{\log_e \pi}\right] = Y \quad , \tag{34}$$

where $D_{M'h}$ is a diagonal matrix of the diagonal elements of $M'h$ and $D_{\log_e \pi}$ is a diagonal matrix of the logarithms of the population proportions. It is clear that the second term on the left hand side of Equation 34 serves merely to add a constant to each column of Y. Therefore if we let

$$Xh = v \tag{35}$$

it is obvious that corresponding columns of Y and v correlate perfectly. Furthermore if we write

$$xh = u \tag{36}$$

and remember that $(I - \frac{1\,1'}{1'\,1})X = x$, we can show that corresponding columns of v and u also correlate perfectly. Now Rao gives for h

$$h = G_w^{-1}M \quad , \tag{37}$$

but it can be proved that

$$m = M(I - D_p\,1\,1') \quad , \tag{38}$$

or

$$M = m + V_M 1' \quad , \tag{39}$$

where V_M is a column vector of the raw score means. Since the origins for the raw scores are entirely arbitrary let us consider two possible sets of solutions for h by Rao's method. The solution based on raw score means is from Equations 37 and 39

$$h_M = G_w^{-1}(m + V_M\,1') \quad . \tag{40}$$

The solution based on deviation score means is

$$h_m = G_w^{-1}\,m \quad . \tag{41}$$

From (40) and (41)

$$h_M = h_m + w\,1' \quad . \tag{42}$$

where

$$w = G_w^{-1} V_M \quad . \tag{43}$$

Consider now

$$Xh_m = u_m \quad , \tag{44}$$

and

$$X(h_m + w\,1') = u_M \quad . \tag{45}$$

From (44) and (45) the covariance matrix between u_m and u_M is

$$C = \frac{u_m' u_M}{N} = h_m' G_T h_m + h_m' G_T w 1' \quad . \tag{46}$$

Suppose now we consider any diagonal element of (46), say i, and dropping the subscript m we write

$$C_{ii} = h_{.i}' G_T h_{.i} + h_{.i}' G_T w \quad . \tag{47}$$

Let Υ be a vector of the diagonal elements of C, and λ a vector of the elements $h_{.i}' G_T h_{.i}$. Then from (47) we can write

$$h_m' G_T w = \Upsilon - \lambda \quad , \tag{48}$$

or from (41) and (43)

$$(m' G_w^{-1})\, G_T\, (G_w^{-1} V_m) = \Upsilon - \lambda \quad . \tag{49}$$

Remember that s is the number of groups and n the number of tests. Let $t = s-1$ if $n \geq s-1$ and $t = n$ if $n < (s-1)$. It can be proved that V_m in (49) can be determined so that t elements of Υ can be chosen at will. The elements in V_m are arbitrary. We can in particular, therefore, choose these elements so that any t elements in Υ vanish. This means that by Rao's method it is possible to choose two different sets of origins for the tests, such that for t of the groups, the classification scores for one solution will have zero correlation with the corresponding classification scores from the other solution. There are an infinite number of such pairs of sets. Although this dependence on origin of test scores might be regarded as a serious defect of Rao's method, it can be shown that the rows of Y in Equation 34 are the same except for a constant if deviation scores rather than raw scores are used. This means that for classification purposes the two sets of origins give the same result.

We rewrite (34)

$$xh_m - \tfrac{1}{2}\,1\,1'\,(D_{m'h} - D_{\log_e \pi}) = y_m \quad . \tag{50}$$

But

$$x = X - 1V_M' \quad , \tag{51}$$

$$m = M - V_M 1' \quad . \tag{52}$$

From (41), (51) and (52) in (50)

$$\left[X - 1V_M'\right] G_w^{-1} \left[M - V_M 1'\right] - \tfrac{1}{2}\,1\,1' \left[D_{(M' - 1V_M')G_w^{-1}(M - V_M 1')} - 2D_{\log_e \pi}\right] = y_m \quad . \tag{53}$$

We let

$$V_M' G_W^{-1} M = v' \quad , \tag{54}$$

$$V_M' G_W^{-1} V_M = k \quad . \tag{55}$$

Expanding (53) and substituting from (54) and (55)

$$X G^{-1} M - X G^{-1} V_M 1' - 1v' + k\, 11' - \tfrac{1}{2}\, 11' \left[D_{M'G^{-1}M} - 2D_{v1'} + k\, I - 2D_{\log_e \pi} \right] = y_m \quad . \tag{56}$$

Now

$$1' D_{v1'} = v' \quad . \tag{57}$$

Let

$$X G^{-1} V_M = \mu \quad . \tag{58}$$

Substituting (37), (57) and (58) in (56)

$$Xh - \mu 1' - \tfrac{1}{2} 11' \left[D_{M'h} - 2D_{\log_e \pi} \right] + \tfrac{1}{2} k\, 11' = y_m \quad . \tag{59}$$

Equation 59 shows that aside from the general constant $\frac{1}{2}k$, the ith row of y_m differs from the ith row in (34) only by the constant μ_i from the vector μ. This means that if a person is assigned to the group for which his classification score in y is highest that the value of μ_i is immaterial.

E. Calculation of the Point Biserial Correlations between Classification Scores and the Dichotomous Criterion Scores

Let

$$xh = y. \quad . \tag{60}$$

We wish to find the point biserial correlations of the y scores with the actual classifications.

Let

$\frac{y'y}{N}$ be the variance-covariance matrix for the y scores,

$\frac{y't}{N}$ the covariance matrix of y with t scores,

$\frac{t't}{N}$ the variance-covariance matrix of t scores.

From (10) and (60)

$$\frac{y'y}{N} = h' G_T h \quad . \tag{61}$$

From (23) and (60)

$$\frac{y't}{N} = h' m D_p \quad . \tag{62}$$

From (19) and the definition of L

$$\frac{t't}{N} = D_p - D_p\, 11'\, D_p \quad . \tag{63}$$

Now let

$D_{h'G_T h}$ be a diagonal matrix of the diagonal elements of $\frac{y'y}{N}$ or the variances of y ,

$D_{(D_p - D_p 11' D_p)}$ be a diagonal matrix of the diagonal elements of $\frac{t't}{N}$ or the variances of t.

By definition, then, the matrix of correlations R_{yt} of the y with the t variables is given by

$$R_{yt} = D_{h'G_Th}^{-\frac{1}{2}} h'mD_p (D_p - D_p^2)^{-\frac{1}{2}}. \tag{65}$$

If we are interested only in the point biserial correlation of each set of classification scores with its corresponding group assignment we may write the row vector of these correlations by

$$1'D_{R_{yt}} = 1'D_{h'G_Th}^{-\frac{1}{2}} D_{h'mD_p} (D_p - D_p^2)^{-\frac{1}{2}}. \tag{66}$$

Now suppose that for any two sets of deviation matrices we have

$$xh - t = \epsilon \quad ,$$

where h is the matrix of least square regression vectors determined so that the trace of $\epsilon' \epsilon$ is a minimum. It is well known that

$$h'G_Th = h' \frac{x't}{N} \quad . \tag{67}$$

Substituting from (23) and (67) in (66) we have, in the case of the least square solution for the h matrix

$$1'D_{R_{yt}} = 1'D_{h'mD_p}^{\frac{1}{2}} \left[D_p - D_p^2 \right]^{-\frac{1}{2}} \quad , \tag{68}$$

or

$$1'D_{R_{yt}} = 1'D_{h'm}^{\frac{1}{2}} \left[I - D_p \right]^{-\frac{1}{2}} \quad . \tag{69}$$

F. Relation of Computational Procedure and Mathematical Development

The steps in the computations for the multiple regression discriminant weights are related to the mathematical development as follows:

Step 1 is simply the matrix M of raw score group means.

Step 2 is the matrix m of deviation score means and is given by Equation 38.

Step 3 is the matrix of deviation means post multiplied by the diagonal matrix of proportions, i.e., mD_p .

Step 4 is given by Equation 10.

Step 5 is given by Equation 24.

Step 6 is given by the term in parenthesis in Equation 28 viz., $(V_p' - V_M'h)$.

The steps in the computation for the point biserial correlations are related to the mathematical development as follows:

Step 7 is simply the product G_Th_w where the subscript w refers to Rao's weights.

Step 8 is the vector of diagonal elements $1'D_{h'G_Th}$ from Equation 61 for Rao's weights.

Step 9 is the vector of diagonal elements $1'(D_p - D_p^2)$ from Equation 63.

Step 10 is the product in row form of $D_{h'G_Th}(D_p - D_p^2)$ where the h's are by Rao's method.

Step 11 is the square root in row form of the elements obtained in step 10, viz., $1' \left[D_{h'G_Th} (D_p - P_p^2) \right]^{\frac{1}{2}}$.

Step 12 is the vector of diagonal elements $1'D_{h_w'mD_p}$ from Equation 62 for Rao's weights.

Step 13 is the vector $1'D_{h_T'mD_p}$ of diagonal elements from Equation 62 for the multiple regression method.

Step 14 gives the vector $1'D_{h_T'mD_p}\left[D_p\,(I - D_p)\right]^{-1}$ for the multiple regression method.

Step 15 gives Equation 66 for Rao's weights.

Step 16 gives Equation 68 for the multiple regression weights.

References

1. Anderson, T. V. Classification by multivariate analysis. *Psychometrika*, 1951, 16 (1), 31-50.
2. Dwyer, P. S. *Linear computations.* New York: Wiley, 1951.
3. Rao, C. R. and Slater, P. Multivariate analysis applied to differences in multivariate groups. *British J. Psychol.*, 1949, 2, 17, Statistics Section.
4. Rao, C. R. Advanced statistical methods in biometric research. New York: Wiley, 1952, Chapter 8.

INTELLIGENCE TEST RESULTS AND OBSERVATIONS OF PERSONALITY DISORDER AMONG 3594 UNWED MOTHERS IN MINNESOTA

JOHN S. PEARSON AND PHYLLIS L. AMACHER

Minnesota Department of Public Welfare
Rochester, Minnesota

Introduction

The zeitgeist of psychological research has witnessed the rise and wane of a number of areas of major emphasis or interest in the first half of the 20th century and an undetermined number of conclusions have entered the body of psychological thinking and knowledge without benefit of anything approximating a crucial test of the original hypothesis.

In the 1920's and 30's, controversy over the nature-nurture question stimulated considerable interest in the intellectual capacity of unwed mothers since that group afforded the basis for study of the influence of environment on the intelligence of large numbers of children reared in foster homes[3, 5, 7, 8]. Since World War II, interest has shifted almost entirely to other areas of investigation leaving a general impression that while unwed mothers might be of almost any degree of intelligence, the total distribution probably averages nearly one standard deviation below the mean for the general population. In a number of studies it has appeared to us that a common sampling error might be operative in that the subjects were usually selected from others known to the agencies by reason of the recorded result of an intelligence test.

At least one author[8] recognized that a bias might thus be introduced through agency policies of testing only those unwed mothers who had not completed high school or otherwise excluding from psychometric testing those where it appeared safe to assume average or better intelligence. On the other hand, several other investigators appear to have overlooked or denied this possibility. In the classic controversy between Skodak[7, 9] and McNemar[4], the contention of the former that ". . . the average IQ of the 80 [tested] mothers was close to that for the total sampling of

The authors wish to express their appreciation to Richard C. Jensen and Harlai R. Levin, University of Minnesota, for major assistance in compiling the data.

154 . . . " was a key point of disagreement with the latter who felt that "... the true fathers were probably above average and that the true mothers were at least average." Skodak had obtained an average IQ of 87.7 for that portion of her sample in which test scores were available.

To the present writers, the opportunity to study individual intelligence test scores for a large sample of Minnesota unwed mothers without any known educational bias appeared to offer something approximating a crucial test of the hypothetical difference between the distribution of intelligence among unwed mothers of children placed in foster homes, and women in general. The importance of the study may well be as much in indicating the need for replication of other investigations of individual differences as in its implications for social agencies dealing with unwed mothers and their children.

Apart from the academic interest in the intellectual abilities of unwed mothers, the group is of importance in that the social and personal problems incident to illegitimate pregnancy are of considerable proportions. Over 80,000 births out of wedlock have been reported annually in the United States in recent years according to the Federal Security Agency. Numerous studies have attempted to analyze emotional and personality disorders which result in or result from illegitimate pregnancy. An example is the report of Kasanin and Handschin [2] on a series of 16 cases of whom seven had had more than one illegitimate child. They concluded that the pregnancies were generally symptomatic of hysterical dissociation states in which the women were acting out Oedipal incestuous phantasies. A number of other studies are similar in treating with small numbers of cases referred or selected because of obvious need for psychiatric treatment and in which psychodynamic mechanisms are emphasized as factors contributing to the illegitimate pregnancy. Schmideberg [6], on the other hand, citing evidence on the rate of premarital sexual experience from the works of Kinsey and Terman represents a smaller group of authors who emphasize situational rather than psychological factors. No previous study however has sought to determine the relative incidence of normal and abnormal personalities in a large sample of unwed mothers.

The data available on our sample of Minnesota mothers permitted only a cursory estimate of the emotional status from social history summaries accompanying referral to the psychologist. Nevertheless, in view of the size of the sample and conditions under which it was obtained, the conclusions drawn may be of some nomothetic value.

Sample

A survey of the records of the Bureau for Psychological Services of the Minnesota Department of Public Welfare revealed that 3594 unwed mothers had been tested for various child placement agencies in this state in the five year period from July 1, 1946 to June 30, 1951. We chose this period for study because the presumptive effects of the special social conditions incident to World War II were thought to have been minimized; because the social history information available for this period was uniform and relatively complete; and because the 1937 revision of the Stanford-Binet was the primary measure of intelligence in use by psychologists of our staff during that time. All examinations were administered and scored by psychologists well trained and highly experienced in the use of this and other tests.

In this five year period, 8627 births out of wedlock were recorded in the state and approximately 3394 illegitimate children were placed for adoption (89% to 90% of all adoptive placements were children of illegitimate parentage). All but one of the 16 licensed adoption agencies in the state utilized the services of the Bureau for Psychological Services, the exception being a Jewish welfare agency handling a relatively small number of cases.

The sample of 3594 cases may be regarded conservatively, as including about 40% of all unwed mothers in Minnesota in a five year period and probably includes more than 95% of such mothers who surrendered their children for adoptive place-

ment by social agencies. Our data imply that about 60% of Minnesota unwed mothers rear their own illegitimate children or place them informally with relatives since state law is quite stringent with respect to foster home or adoptive placement, and "black market" placement of illegitimate babies in the period covered was virtually non-existent. Agency practice by the Child Welfare Unit of the Minnesota Department of Public Welfare is to require of all public and private placement agencies under its supervision that case information include results of an individual intelligence test of the unwed mother wishing to place her child even when evidence such as a college or graduate degree would appear to rule out the possibility of intellectual deficiency. Unfortunately for the purposes of this study, fathers are rarely examined even when paternity has been established.

The Stanford-Binet Form L was utilized in approximately 86% of the 3594 examinations with the remainder divided about equally between Form M and the Wechsler-Bellevue Scale Form I.

Results

The mean IQ for the total sample of 3594 was 100.19 with a standard deviation of 18.36. The age range in this group was from 13 to 47 years with a median age of 20.64 years. Changing social and educational influences between 1937 and the period covered by the study plus the fact that the standardization group for the Stanford-Binet did not include any individuals beyond 18 years of age make it difficult to assume any particular parameter values for the general population of women in this age range even if one considers that the influence of scores on the Wechsler-Bellevue in our sample may properly be ignored. Nevertheless, in the pragmatic sense, the obtained values for average IQ and variability fall close to an assumed mean of 100.00 and sigma of 16.00 which authorities might agree would represent a good guess.

A further investigation of the intelligence of unwed mothers was undertaken by eliminating from consideration all persons who took tests other than the Stanford Form L and in addition, eliminating Negroes and American Indians regardless of test utilized. A sample of 2975 cases remained with an average IQ of 100.53 and standard deviation of 18.00. Again, the differences from any parameters one might logically assume can hardly be regarded as startling. Fitting a normal curve to this data and computing the X^2 values for differences between observed and theoretical frequencies by .5 S. D. intervals yielded a value of 31.36 which with 13 degrees of freedom falls between the .01 and .001 significance levels. This would indicate that the observed distribution is significantly different from the theoretical Poisson distribution. The explanation for this difference may well be sociological: It was produced chiefly by a notable excess of observed cases in the defective range (IQ 50 to 71) and in the bright normal to superior range (IQ 119 to 127) together with a deficiency in the number of cases observed in the dull normal range (IQ 83 to 91). This suggests that unmarried mothers of defective mentality may come to the attention of child placement agencies by reason of their obvious inability to manage and care for an illegitimate child and that those of bright normal intelligence frequently place their children for adoption through fear of stigma and social ostracism while those of dull normal ability more frequently keep and rear their illegitimate babies because their social milieu is more accepting of such arrangements.

A further analysis of the sample of 2975 cases was made in comparing those 2747 women who up to the time of examination were known to have had only one illegitimate pregnancy with the 228 who had had two or more (the maximum encountered in this study was seven). The mean IQ for the "novices" was 101.15 with a standard deviation of 18.22. The median was 100.84. For the "repeaters," the mean was 93.29; the standard deviation was 17.40; and the median was 92.50. These distributions are presented graphically in Figure 1 which also illustrates the rather close approximation of a normal curve by the combined distribution. From these results it appears safe to conclude that women who become illegitimately pregnant two or more times are likely to be somewhat lower in intelligence than those who give indi-

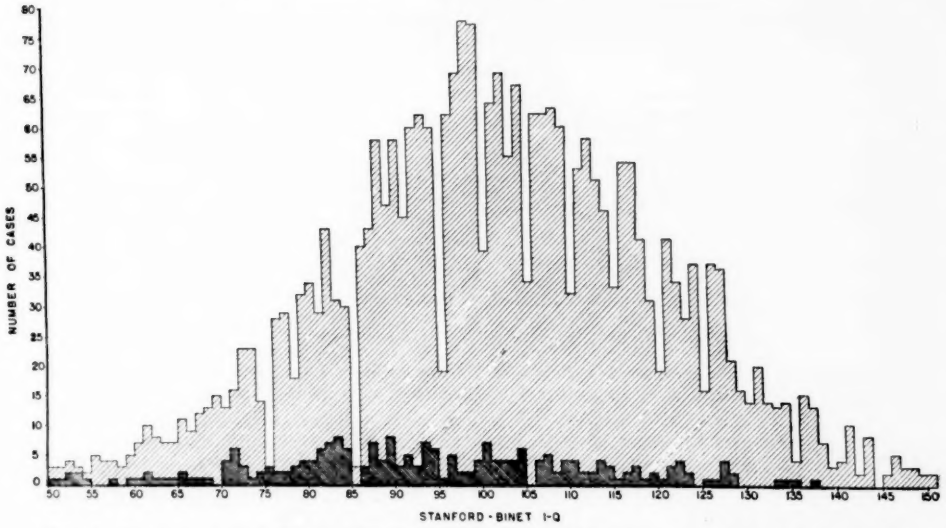

FIG. 1. DISTRIBUTION OF IQ'S OF 2975 UNWED MOTHERS IN MINNESOTA, LIGHTER AREA INDICATES MOTHERS WITH ONE ILLEGITIMATE CHILD (N = 2747); DARKER AREA INDICATES THOSE WITH TWO OR MORE (N = 228)

cation of profiting from their first experiences. It is noteworthy, however, that the average IQ of our "repeaters" was higher than that generally reported elsewhere for unwed mothers in general.

We attempted to estimate the incidence of personality disorder among our subjects from review of the social caseworker's reports accompanying the referrals to the psychologist. These varied in length from a few sentences to several typewritten pages in the 3450 cases where we felt there was at least some basis for judgment. Only evidences of social adjustment or lack of it prior to the illegitimate pregnancy were considered since we were interested in personality disorder as a causal factor rather than as a resultant among unwed mothers. A decision was made by either of us as to whether or not emotional maladjustment was present in significant degree prior to the illegitimate pregnancy. We felt on the basis of joint evaluation of the first 100 case histories that our agreement in rating was sufficiently high to permit division of labor on the remainder of the cases since the validity of the whole procedure would remain in doubt regardless of the reliability of ratings. Cases were categorized as "normal personality", "neurotic", "psychotic", "sociopathic" or "psychopathic", "primary mental deficiency", and "miscellaneous neurological disorders". Obviously, these judgments were highly subjective particularly as regards the individuals judged to be normal or neurotic. The persons included in the "miscellaneous neurological disorders" group included several cases of cerebral palsy, epilepsy, and birth injury which were presumed to have an important influence on personality development or otherwise to have a direct bearing on premarital sexual activity. We considered separately those cases where mental deficiency appeared to be of primary importance as a causative influence in the illegitimate sexual activity and these were not otherwise categorized although a number gave obvious indications of serious emotional maladjustment of one kind or another.

The results of this phase of the investigation are summarized in Table 1. Here it may be seen that 2506 or 72.6% of the 3450 unwed mothers were judged to be normal with respect to emotional and behavioral status. Again, no estimates are available as to the relative frequencies with which women in the general population would be

placed in our different categories according to judgments of social histories equivalent to those we surveyed; however, the incidence of psychosis and mental deficiency by our standards would appear to be in rather close approximation to the corresponding estimates for women in this age range in the general population. In general, we tended to follow a rather conservative course in categorizing cases as abnormal personalities only when there was definite behavioral evidence which we regarded as indicative of psychopathology. In the absence of any reliable sampling of women selected at random from the general population, we can only hope that the incidence of rather serious personality or behavioral inadequacy would not approach the 27.4% which we encountered in our sample of unwed mothers.

TABLE 1. JUDGMENTS OF PERSONALITY OR BEHAVIORAL STATUS OF 3450 MINNESOTA UNWED MOTHERS GROUPED BY INTELLECTUAL LEVEL.

	Behavioral or Personality Evaluation.						
Intellectual Level	Normal	Neurotic	Psychotic	Psychopathic (sociopathic)	Misc. Neurological Disorders	Primary Mental Deficiency	Totals
Low (IQ<90)	675	208	21	41	10	116	1071
Average (IQ 90-109)	1050	279	3	42	2		1376
High (IQ 110>)	781	170	4	40	8		1003
Totals	2506	657	28	123	20	116	3450

Tests of significance of differences in frequency proportions revealed that the overall incidence of personality deviation appeared to be higher among unwed mothers of low intelligence than among those of average or high ability, while the latter groups did not differ significantly from each other. The reason for this might lie in a general negative correlation between emotional adjustment and intelligence as postulated by Dewan from his study of 30,000 Canadian servicemen[1]. On the other hand, the social workers might have tended to repress evidence of instability in their reports concerning the more intelligent unwed mothers because of unconscious identification or some other factor. Certainly in many instances it was evident from the reports we reviewed that social workers found difficulty in empathizing with the less intelligent clients. An example is the case of one unpersonable woman of IQ 82 then expecting her third illegitimate child. The worker scarcely concealed distaste as she wrote among other things, "She doesn't feel she has been promiscuous as she always said, 'No'. But they went right ahead anyway."

SUMMARY

This investigation was concerned with distribution of intelligence and incidence of personality or behavior disorders among 3594 Minnesota unwed mothers. The sample comprised about 40% of all unwed mothers in the state over a five year period and is believed to include more than 95% of those who placed their illegitimate children for adoption. Results of individual intelligence tests indicated an average IQ for the whole group of 100.19 with a standard deviation of 18.36. The distribution differed significantly from the normal curve by reason of an excess of cases observed in the mentally deficient and the bright normal to superior ranges together with a deficiency in number of cases observed in the dull normal range. The writers believe that the explanation for this difference is sociological relating to foster home placement and that the distribution of intelligence among unwed mothers in general is a very close approximation of the distribution for women in general. This conclusion is in contrast to findings in a majority of earlier reports.

Evidence of personality or behavior disorder antedating the illegitimate pregnancy was judged to be present in 27.4% of the cases according to a rather conservative interpretation of social history information available in 3450 cases. Personality and behavior disorder appeared to be somewhat more frequent among unwed mothers of low intelligence than among those of average or high intellectual ability.

References

1. Dewan, J. Intelligence and Emotional Stability. *Am. J. Psychiat.* 104, 1948, 548 ff.
2. Kasanin, J. and Handschin, S. Psychodynamic Factors in Illegitimacy. *Am. J. Orthopsychiat.*, XI, 1, 1948.
3. McClure, W. Intelligence of Unmarried Mothers. *Psychol. Clin.*, 20, 1931, 154-157.
4. McNemar, Q. Critical Examination of the University of Iowa Studies of Environmental Influences upon the IQ. *Psychol. Bull.*, 37, 2, 1940, 74-79.
5. Nottingham, R. A Psychological Study of Forty Unmarried Mothers. *Genet. Psycho. Mono.* 19, 1937, 155-228.
6. Schmideberg, M. Psychiatric Social Factors in Young Unmarried Mothers. *Bull. National Comm. on Service to Unmarried Parents*, Cleveland, Ohio, Oct. 1952.
7. Skodak, M. Children in Foster Homes: A Study of Mental Development. *Univ. Ia. Stud. Child Welf.*, 16, 1, 1939.
8. Snygg, D. The Relation Between the Intelligence of Mothers and of their Children Living in Foster Homes. *J. Genet. Psychol.*, 52, 1938, 401-406.
9. Wellman, B., Skeels, H. and Skodak, M. Review of McNemar's Critical Examination of Iowa Studies. *Psychol. Bull.*, 37, 2, 1940, 96-106.

FACTORS OF IMPORTANCE FOR CREATIVITY

JOHN E. DREVDAHL

VA Hospital, Minneapolis
and
University of Minnesota Medical School

Problem

Despite the fact that creativity has been a matter of interest among psychologists for many years, most studies and discussions of the subject have been primarily speculative in nature. There have been three general approaches to the study of creativity: the historical-anecdotal approach, best illustrated by Kretschmer[6], the introspective personal report approach exemplified in Patrick's studies[7, 8], and the test approach, used by Terman[14] and Roe[13, 11]. With the exception of less than a score of relatively recent studies, some reported and some still under way, the experimental or objective (test) approach to the study of creativity has been studiously neglected. The purpose of this study was to explore some of the possible relationships between ratings of creativity in a high level population and certain objectively measured personality and intellectual factors.

Procedure

The subjects used in this study were graduate or advanced undergraduate students from several of the science and arts departments of the University of Nebraska. All of the subjects were under the close individual supervision of the raters in their area of specialization at the time of testing and for some time preceding it. The raters were faculty members of the University of Nebraska experienced in the task of evaluating the creative abilities of their students.

In order to assure, in so far as possible, the relative independence of the intellectual and personality characteristics under investigation, several factor tests were selected and administered. The personality characteristics were measured by means of Cattell's Sixteen Personality Factor Questionnaire[1]. The intellectual characteristics were measured by means of Thurstone's Primary Mental Abilities

Test[15]. In view of Guilford's[3] suggestion that the standardized and accepted intelligence tests may not include many of the critical variables that may be involved in creative effort, a group of tests was selected from his Factor Analytic Study of Creative Thinking[4]. Guilford found, in addition to the expected intellectual factors, nine other interpretable factors that may be involved in creative thinking.[1] For the purposes of this study, the test with the highest loading on each of the factors was used as a measure of that factor.[2]

The tests were administered to the subjects in a consistent order during two testing sessions. Of 90 subjects, 76 completed the first session (tests of the intellectual factors) and 64 completed both sessions. Several subjects were dropped because the raters felt that they did not know them well enough to be able to make a valid rating, and several (foreign students) were dropped because of language handicaps. As a result of these omissions, the final group used in the analysis of the intellectual factors numbered 64, while the group used in the analysis of the personality factors numbered 52.

Two raters rated each subject independently and made two ratings on each subject, using each time a seven point scale of creativity. The first rating was based upon the rater's own definition of creativity, and the second rating, which took place at least two weeks after the first, in order to minimize the carry-over effects from one set of ratings to the other, was based upon a given definition—a distillation of the various definitions that have been proposed in the literature. There were no significant differences between the defined and undefined ratings, and the given definition may be considered to include the various subjective definitions of the raters. The given definition is as follows:

> "Creativity is the capacity of persons to produce compositions, products, or ideas of any sort which are essentially new or novel, and previously unknown to the producer. It can be imaginative activity, or thought synthesis, where the product is not a mere summation. It may involve the forming of new patterns and combinations of information derived from past experience, and the transplanting of old relationships to new situations and may involve the generation of new correlates. It must be purposeful or goal directed, not mere idle fantasy—although, it need not have immediate practical application or be a perfect and complete product. It may take the form of an artistic, literary or scientific production or may be of a procedural or methodological nature."

The subjects were divided into creative and non-creative groups and into arts and science groups. Those subjects whose creativity ratings fell below the fiftieth percentile for the whole group were placed in the non-creative group. The correlation (Pearson) between different judges' ratings on all subjects was .71. The significance of the differences obtained between the various groups was determined by means of a 2 x 2 analysis of variance on each of the 30 factors under investigation. For each of the variables on which the groups showed or closely approached a significant difference, a biserial correlation coefficient was calculated in order to determine the degree of relationship between the scores on that variable and the creativity classification.

Results

Intellectual Factors

As can be seen in Table 1, there were no significant differences found between the creative and non-creative groups on a one-test estimate of Guilford's factors called redefinition, closure, ideational fluency, spontaneous flexibility, associational fluency, or sensitivity to problems. There was a tendency, although not statistically

[1]The writer wishes to express his appreciation to Dr. J. P. Guilford for making his tests available for experimental purposes. This study was done at the University of Nebraska.

[2]For reasons of practicality the test with the second highest loading on "adaptive flexibility" was used as a measure of this factor.

TABLE 1. ANALYSIS OF VARIANCE RATIOS AND SIGNIFICANCE LEVELS FOR CREATIVE VS. NON-CREATIVE (A) AND ARTS VS. SCIENCE (B) GROUPS ON ALL FACTORS

Factor	Source of Variance	F	Factor	Source of Variance	F
Word Fluency (Guilford)	A	3.52	General Intelligence *vs.*	A	<1.00
	B	<1.00	Mental Defect	B	<1.00
Redefinition	A	1.65	Ego Strength *vs.*	A	1.29
	B	5.33*	General Neuroticism	B	<1.00
Adaptive Flexibility	A	3.62	Ascendance *vs.* Submission	A	<1.00
	B	<1.00		B	<1.00
Associational Fluency	A	2.22	Surgency *vs.* Desurgency	A	8.35**
	B	<1.00		B	<1.00
Closure	A	1.46	Super Ego Strength *vs.*	A	<1.00
	B	<1.00	Lack of Internal Standards	B	<1.00
Originality	A	4.76*	Adventurous Cyclothymia	A	<1.00
	B	<1.00	*vs.* Withdrawn Schizothymia	B	3.28
Ideational Fluency	A	<1.00			
	B	<1.00	Emotional Sensitivity *vs.*	A	<1.00
Sensitivity to Problems	A	1.52	Tough Maturity	B	23.27**
	B	1.05	Paranoid Schizothymia *vs.*	A	<1.00†
Spontaneous Flexibility	A	<1.00	Truthful Accessibility	B	<1.00
	B	<1.00	Bohemianism *vs.*	A	<1.00
Verbal Meaning	A	5.82*	Practical Concernedness	B	21.52**
	B	<1.00	Sophistication *vs.*	A	2.13
Space	A	<1.00	Rough Simplicity	B	<1.00
	B	<1.00	Anxious Insecurity *vs.*	A	<1.00
Reasoning	A	<1.00	Self Confidence	B	2.52
	B	<1.00	Radicalism *vs.*	A	5.09*†
Number	A	<1.00	Conservatism	B	<1.00
	B	9.40**†	Self Sufficiency *vs.*	A	4.21*†
Word Fluency (Thurstone)	A	<1.00	Lack of Resolution	B	<1.00
	B	<1.00	Will Control and Stability	A	<1.00
General Intelligence	A	<1.00		B	3.23
	B	2.25	Nervous Tension	A	<1.00
Cyclothymia *vs.* Schizothymia	A	4.53*		B	1.68
	B	<1.00			

Note: The probability of obtaining ten statistics significant at the .05 level for 62 calculated statistics is approximately .001 [12].

*Significant beyond the .05 level of confidence.

**Significant beyond the .01 level of confidence.

†Significant interaction. See text for explanation.

significant, for the creative group to score higher on this latter factor. The creative group scored significantly higher than the non-creative group on the factor of originality, and although the difference between the groups on the factors of word fluency and adaptive flexibility was not statistically significant at the .05 level, as can be seen in Table 2, there appears to be a low, although significant, relationship between

TABLE 2. BISERIAL CORRELATIONS BETWEEN CREATIVITY CLASSIFICATION AND FACTOR SCORES FOR ALL SUBJECTS

Factor	r_{bis}
Word Fluency	.37**
Originality	.33**
Adaptive Flexibility	.29*
Verbal Meaning	.37**
Cyclothymia *vs.* Schizothymia	.36*
Surgency *vs.* Desurgency	.47**
Radicalism *vs.* Conservatism	.38**
Self Sufficiency *vs.* Lack of Resolution	.33

* Significant beyond the .05 level.
** Significant beyond the .01 level.

creativity and higher scores on these two factors. The science group scored significantly higher than did the arts group on the factor of redefinition, but there were no other significant differences between the arts and science groups on Guilford's factors. Inasmuch as redefining activity would seem to be common to both artists and scientists, some explanation might be considered at this point. Redefinition of the type required by this test involves the use of substitute objects for various purposes. This specific activity is quite common in the science laboratory where the construction of temporary and makeshift equipment is a continual necessity. On the other hand, artistic redefinition may involve, for the most part, an interpretative redefinition of an existent work without involving actual structural substitutions of the type required in the laboratory. This type of artistic redefinition is poorly measured, if at all, by the test used in this instance.

Table 1 also shows that the creative group scored significantly higher than the non-creative group on the factor of verbal meaning on Thurstone's Primary Mental Abilities Test. There were no other significant differences between the creative and non-creative groups on this test. There was a significant difference between the arts and science groups on the number factor, but also a significant interaction. Inspection of the data suggests that this difference is primarily attributable to the exceptionally high scores of the creative science group. Numerical facility (number) may be merely a symbolic counterpart of verbal facility, and inasmuch as the transformation of concepts and relationships, ordinarily dealt with in verbal symbols, into mathematical symbols, is more characteristic of the sciences than the arts, it might be expected that creative persons in science would display greater facility in handling such concepts than would creative persons in fields where verbal facility is sufficient. In view of this, perhaps number facility may be considered simply a special instance of verbal facility (word fluency, verbal meaning).

Some comment concerning the meanings of these factors and their interrelationships might justifiably be considered at this point. Word fluency and verbal meaning denote a wide knowledge of things, ideas, and concepts expressed in words. Adaptive flexibility is a characteristic that involves the ability to change set or approach to meet new requirements imposed by changing conditions or problems. It is a more restricted type of flexibility than most (e.g.: spontaneous flexibility), inasmuch as the direction of change is dictated by the situation or problem demanding solution. Associational fluency involves the ability to make relevant associations quickly and easily. It is somewhat similar to the preceding factor in its demand for relevancy. Originality involves the ability to produce uncommon, unusual, and "clever" responses. We might, therefore, define creative thought, in terms of the present findings, as the goal directed, easily flexible, manipulation of knowledge (concepts, relationships, etc.), in a wide variety of novel or original ways. Such a process fulfills the specifications for creativity as proposed by Spearman [13], Ribot [9], or Hutchinson [5], or the specifications required by the given definition used in this study. The knowledge referred to here (arising from word fluency and verbal meaning) can be

fostered by known means of education so what remains of critical importance appears to be the less educable factors of adaptive flexibility, associational fluency, and perhaps of greatest importance, originality. Evaluation of these three intellectual factors might be of considerable importance in the selection of potentially creative persons if these results can be substantiated in further investigations.

Although there does not appear to be a significant relationship between creativity and general intelligence in this population,[3] it would be erroneous to assume that general intelligence or the other nonsignificant intellectual factors studied are unimportant. It seems more likely that these factors may be necessary but insufficient characteristics for creativity. Also of interest is the finding that, for the most part, the creative artist possesses the same intellectual characteristics as the creative scientist.

Personality Factors

Table 1 shows that there were no significant differences between the creative and non-creative groups on the factors of intelligence *vs.* mental defect; ego strength *vs.* general neuroticism; ascendance *vs.* submission; super-ego strength *vs.* lack of internal standards; adventurous cyclothymia *vs.* withdrawn schizothymia; emotional sensitivity *vs.* maturity; bohemianism *vs.* practical concernedness; sophistication *vs.* simplicity; anxious insecurity *vs.* self confidence, will control and stability, or nervous tension. The creative group scored higher than the non-creative group on the factors of radicalism *vs.* conservatism,[4] and self sufficiency *vs.* lack of resolution; and lower on the factors of cyclothymia *vs.* schizothymia, and surgency *vs.* desurgency. On the factors of radicalism *vs.* conservatism, and self sufficiency *vs.* lack of resolution, there was a significant interaction. Inspection of the data suggests that the creative art group was significantly more radical and self sufficient than the creative science group. In terms of these results we might characterize the creative group as somewhat withdrawn and quiescent, more concerned with ideas and things than with people. Creative artists we may further characterize as considerably more radical and self sufficient. These latter characteristics may reflect the creative artists' tendency to break away from the routinized and accepted patterns of work and behavior. A further question of interest concerns the creative individual's preoccupation with his work. Does this occur because the individual is unable to find gratifications in his relationships with people, or do they simply find sufficient gratification in their work and, therefore, have less need for interpersonal relationships?

There were also a number of significant differences between the art and science groups apart from their classification as creative or non-creative (Table 1). The art group scored significantly higher than did the science group on the factors of emotional sensitivity *vs.* maturity, and bohemianism *vs.* practical concernedness; and the science group scored lower, although not significantly so, on the factors of adventurous cyclothymia *vs.* withdrawn schizothymia and nervous tension. There was a significant interaction on the factor of paranoid schizothymia *vs.* truthful accessibility, and the data indicate that the creative art and non-creative science groups scored high on this factor. Thus, it appears that suspiciousness and interest in the internal mental life, or possibly interoceptiveness, may be associated with creativity in the arts, but that trusting attitudes and interest in external things, or extroceptiveness, may be more associated with creativity in the sciences.

Discussion

Inspection of the original data suggests that the differences found in this study may be minimal estimates of true differences. The use of more highly loaded tests, or of subjects whose creativity has been demonstrated, might magnify the differences found here, as might the use of the upper and lower thirds rather than the upper and lower halves on the creativity ratings. It might be reasonable to expect,

[3] This does not seem to be completely attributable to the intellectual homogeneity of the group inasmuch as there was a range of over 50 IQ points within the total group

[4] When factors are listed in bipolar terms, high scores indicate predominance of the first named characteristic.

for example, that further refinement of techniques may demonstrate the existence of the factor "Sensitivity to Problems" that is shown to be of only doubtful significance in this study.

The results of this study provide considerable support for the contention that individuality and non-conformity are desirable for creative effort. Several of the intellectual factors are, to a great extent, measures of the degree to which the individual can avoid the usual, routine, conventional ways of doing and thinking of things. As one might expect, this ability, while common to all creative persons, may vary in its degrees of significance and inclusiveness according to its area of application. As far as the science fields are concerned, it may be necessary for the creative individual to be conventional in accepting knowledge, known methods and, perhaps, goals. Once these are accepted and internalized, the scientist's unconventionality (which is more intellectual than personal) can be expressed within this framework by means of his flexibility and fluency in approach, and in his originality. With respect to the art fields, on the other hand, the creative artist may be less dependent upon what has gone before, and may depend to a greater extent upon his personal unconventionality.

A further investigation of the personality factors associated with creativity in science, in a group whose creativity has been conclusively demonstrated, recently has been completed by Cattell and Drevdahl[2], and provides affirmation of some of the findings of this study. A similar study of creativity in the arts is under way.

Summary

An investigation was made into some of the relationships between certain objectively measured intellectual and personality characteristics and ratings of creativity in a population of advanced undergraduate and graduate students at the University of Nebraska. The results of this study suggest the following conclusions:

1. Creative persons in this group appear to be superior to non-creative persons in their verbal facility, fluency, flexibility, and in their originality.
2. Creative persons in this group appear to be considerably more withdrawn and quiescent than non-creative persons. Creative artists were somewhat more radical and self sufficient than creative scientists or non-creative persons in either the sciences or the arts.
3. Apart from its classification as creative or non-creative, the art group was more sensitive emotionally, and more bohemian.
4. Individuality or non-conformity appear to be desirable for creativity.

References

1. Cattell, R. B., Saunders, D. R. and Stice, G. *Handbook for the sixteen personality factor questionnaire.* Champaign, Ill.: Institute for Personality and Ability Testing, 1952.
2. Cattell, R. B., and Drevdahl, J. E. *A comparison of the personality profiles of eminent researchers with those of eminent teachers and administrators and the general population. British J. Psychol.*, in press.
3. Guilford, J. P. Creativity. *American Psychologist*, 1950, 5, 444-454.
4. Guilford, J. P., Wilson, R. C., and Christensen, P. R. *A factor analytic study of creative thinking: II Administration of tests and analysis of results.* Reports from the Psychological Laboratory, #8. Los Angeles: Univ. of So. Calif., 1952.
5. Hutchinson, E. D. *How to think creatively.* New York: Abingdon-Cokesbury Press, 1949.
6. Kretschmer, E. *The psychology of men of genius.* London: K. Paul, Trench, Trubner & Co. Ltd., 1931, translated and introduced by R. B. Cattell.
7. Patrick, Catherine. Creative thought in poets. *Archives of Psychology*, 1935, 26, 5-74.
8. Patrick, Catherine. Scientific thought. *J. Psychol.*, 1938, 5, 55-83.
9. Ribot, T. *Essay on the creative imagination.* Chicago: Open Court Publishing Co., 1906.
10. Roe, A. The personality of artists. *Educ. Psychol Measmt.*, 1946, 6, 401-8.
11. Roe, A. Analysis of group Rorschachs of biologists. *J. Proj. Tech.*, 1949, 13, 25-43.
12. Sakoda, J. M., Cohen, B. H., and Beall, G. Test of significance for a series of statistical tests. *Psychol. Bull.*, 1954, 51, 172-5.
13. Spearman, C. E. *Creative mind.* New York: D. Appleton & Co., 1931.
14. Terman, L. M. Scientists and nonscientists in a group of 800 gifted men. *Psychol. Monogr.*, 1954, 68, 7.
15. Thurstone, L. I. and Thurstone, T. G. *Examiner manual for the s. r. a. primary mental abilities.* Chicago: Science Research Associates Inc., 1949.

THE OBJECTIVE RORSCHACH:

A SUGGESTED MODIFICATION OF RORSCHACH TECHNIQUE

P.O. O'REILLY, M.B., B.CH.**

Saskatchewan Hospital, North Battleford, Saskatchewan Canada

PROBLEM

In recent years several attempts[1-6] have been made to adapt the Rorschach Test for screening purposes with emphasis on speed and objectivity of scoring of responses for the differentiation of normals from neurotics on a ranking basis. This study describes an objective method based on a modification of the Rorschach leading to differentiation of normals, neurotics and psychotics for research purposes primarily.

CONSTRUCTION OF SCALE

A preliminary scale was set up with twelve possible choices for each card by scanning previous definitely "neurotic" and "psychotic" Rorschach records of patients admitted to the Saskatchewan Hospital, North Battleford. Four typically "neurotic" and four typically "psychotic" responses were selected for each card. 100 neurotic and 100 psychotic responses were scanned and frequency of occurrence decided selection. As a basis for the normal responses, 50 student nurses who had not as yet been inculcated with the sciences of nursing were asked to make two responses to each card and four typically normal responses were chosen on the basis of frequency of occurrence from the 100 responses made to each card. A preliminary scale of 12 choices for each card was administered to 50 other student nurses who were asked to rank the responses in order of frequency. Their preferential ranking of responses fell into a hierarchy from normal to neurotic to psychotic, so the preliminary scale was retained for further study. The nature of the material used is illustrated in Table 1 which comprises the twelve selective responses proposed for each of the 10 cards.

TABLE 1. SELECTIVE RESPONSES FOR EACH RORSCHACH CARD
(No = Normal; Ne = Neurotic; Ps = Psychotic)

CARD ONE

(Ne)	A dark cloud	( )
(No)	A pelvis	( )
(Ne)	Silhouetted figure of a woman	( )
(Ps)	Triangles (white spaces)	( )
(No)	A bat	( )
(Ps)	A clammy toad	( )
(Ne)	An x-ray	( )
(No)	A map	( )
(Ps)	The back of a dog	( )
(No)	Two people dancing	( )
(Ne)	A black, ugly bat	( )
(Ps)	A wild cat	( )

*Saskatchewan Committee on Schizophrenia Research; supported by the Department of National Health and Welfare, Canada.

**I wish to acknowledge the valuable assistance and helpful suggestions of the following people who were associated in this Research Project: Duncan Blewett, M. A., Ph.D , Consulting Psychologist for Saskatchewan, B. Prodanchuk, B. A., Research Psychologist, K. McLaren, B. A. M. Ed., Director of Training and Dr. G. Fisher and Dr. M. Demay, Superintendent, Saskatchewan Hospital, North Battleford, who provided the facilities, assisted in drawing up the technique of the project and provided helpful comments and suggestions.

CARD TWO

(No) Two bears.....( )
(Ne) Black paint spattered, with red underneath.....( )
(Ne) Blood spots.....( )
(No) Two people.....( )
(Ps) A gladiola.....( )
(No) Two dogs.....( )
(Ps) A triangle (centre space).....( )
(Ne) X-ray of part of the body.....( )
(Ps) The devil.....( )
(No) Two children playing pat-a-cake.....( )
(Ps) The heart taken out of somebody's body.....( )
(Ne) A cloud.....( )

CARD THREE

(No) Two men fighting.....( )
(Ne) Blood marks.....( )
(Ps) A shrimp.....( )
(Ne) A shady lane with trees on either side.....( )
(Ne) An x-ray of the spine.....( )
(No) A red bow tie.....( )
(Ps) False teeth.....( )
(Ne) Shadows of two people.....( )
(No) A butterfly.....( )
(Ps) An actor's face.....( )
(No) A pelvis.....( )
(Ps) An x-ray of the head.....( )

CARD FOUR

(Ps) A Ukrainian devil, and lots of dead people.....( )
(Ne) Smoke.....( )
(Ne) The lungs.....( )
(No) An animal khid or skin.....( )
(Ps) The Black Death.....( )
(No) Man or ape sitting or lying.....( )
(Ps) A diseased person.....( )
(No) An animal's skull.....( )
(Ne) Darkness.....( )
(Ps) Half a bear walking.....( )
(Ne) An x-ray.....( )
(No) Boots.....( )

CARD FIVE

(Ps) A rabbit with wings.....( )
(No) A butterfly.....( )
(Ne) Smoke.....( )
(Ps) Two men with their heads joined together.....( )
(Ne) An aerial view of an island.....( )
(No) A rabbit.....( )
(Ps) An eagle flying standing on the ground.....( )
(Ne) A dark storm cloud.....( )
(No) A bird.....( )
(Ne) A horrible looking black bat.....( )
(Ps) A snail that is clammy and cold.....( )
(No) A bat.....( )

CARD SIX

(Ne) The human body showing the backbone and the back.....( )
(No) An insect.....( )
(Ps) The legs of a chair.....( )
(No) A diamond drill.....( )
(Ne) An x-ray.....( )
(Ne) Scenery reflected in water.....( )
(No) A fountain.....( )
(Ps) A map; turtle drinking water.....( )
(Ne) An aerial view of a river with steep banks.....( )
(No) An animal skin.....( )
(Ps) Flash bulbs with blue points.....( )
(Ps) The interior of a water melon.....( )

CARD SEVEN

(Ps) A rubber scraper.....( )
(No) Two girls dancing.....( )
(Ne) Smoke.....( )
(Ne) A dark piece of land with water around it as seen from a plane...( )
(No) Two rabbits.....( )
(Ps) Radar.....( )
(Ne) A map of an island.....( )
(Ne) A road and a house seen in the distance.....( )
(Ps) A broken heart.....( )
(No) Two ladies talking.....( )
(Ps) An army tank with a gun in front.....( )
(Ne) Clouds.....( )

CARD EIGHT

(Ne) Parts of the body.....( )
(No) A coat of arms.....( )
(Ps) The interior of an orange.....( )
(No) Flowers.....( )
(Ne) A drawing in a medical book.....( )
(Ps) Heavenly clouds...dark nebulae.....( )
(No) A skeleton.....( )
(No) Two animals.....( )
(Ps) A ship.....( )
(Ne) Scenery reflected in water.....( )
(Ps) A spider.....( )
(Ne) A snow covered mountain in the distance.....( )

CARD NINE

(No) Sea horses.....( )
(Ne) A diagram of the organs of the body.....( )
(No) A painting.....( )
(Ps) A ballet ensemble.....( )
(Ne) A drawing from a hygiene book in school.....( )
(No) A crab.....( )
(Ne) A design on material.....( )
(No) Two men.....( )
(Ps) Blood poisoning.....( )
(Ne) Dark clouds with the sun shining on them.....( )
(Ps) Two people with tusks on them.....( )
(Ps) A monkey.....( )

CARD TEN

(No) Snakes, mice, rats.....( )
(Ps) Water running in all directions.....( )
(Ne) Some drawing from a doctor book.....( )
(Ps) Scribbling.....( )
(No) Two men.....( )
(Ps) A sweet pea connected to a sea horse.....( )
(No) Bugs and beetles.....( )
(Ne) Fire and smoke.....( )
(Ne) Various internal organs of the body.....( )
(No) A deep sea scene.....( )
(Ps) A rabbit's head with worms coming out of the eyes.....( )
(Ne) Two darkies looking at each other.....( )

Procedure

The objective Rorschach was administered individually to each subject and instructions were given as follows: "You are going to be shown a set of 10 cards with designs on them. For each card on this sheet of paper there are 12 answers. Read the listed answers and choose, if possible, two answers which you think best describe the card and place a check in the brackets. However, if you cannot see any of these answers, there is no necessity to mark any."

A total score was obtained by arbitrarily assigning numerical scores of 3 for psychotic choices or rejections (no responses), 2 for neurotic choices, and 1 for normal choices. The reliability of this objective Rorschach scoring was determined by readministering the test six weeks later to 50 student nurses, with the obtained $r = .91$ indicating satisfactory reliability.

The results to be reported in table 2 were obtained by comparing three groups including 50 normals, 30 neurotics and 80 psychotic patients. The characteristics of these groups are given below:

Group I consisted of 50 normals composed of nurses, stenographers, clerks and administrative staff members of the Saskatchewan Hospital, North Battleford, chosen at random from an occupational listing.

Group II consisted of 30 neurotics, 90% of whom were recent admissions and who had been diagnosed by staff conference consisting of medical superintendent, clinical director and staff psychiatrists, nine in number, as psychoneurotics of various types.

Group III consisted of 80 patients, diagnosed by staff conferences as psychotic, 85% of whom were recent admissions.

Results

Table 2 presents the scores obtained by normal, neurotic and psychotic groups.

Table 2. Scores Obtained by Normal, Neurotic and Psychotic Groups

Scores	Normals (N–50)			Neurotics (N–30)			Psychotics (N–80)		
	f	%	Cum.%	*f*	%	Cum. %	*f*	%	Cum. %
24	2	4	4						
25	2	4	8						
26	1	2	10						
27	6	12	22						
28	7	14	36						
29	7	14	50						
30	10	20	70						
31	4	8	78	3	10.0	10.0			
32	3	6	84	4	13.3	23.3			
33	7	14	98	5	16.6	39.9			
34				2	6.6	46.5	2	2.5	2.5
35				4	13.3	59.8			
36	1	2	100	6	20.0	79.8			
37				5	16.6	96.4	14	17.5	20.0
38				1	3.3	99.7	2	2.5	22.5
39							12	15.0	37.5
40							2	2.5	40.0
41							6	7.5	47.5
42							1	1.2	48.7
43							4	5.0	53.7
44							7	8.7	62.4
45							1	1.2	63.6
46							2	2.5	66.1
47							4	5.0	71.1
48							4	5.0	76.1
49							2	2.5	78.6
50							1	1.2	79.8
51									
52							1	1.2	81.0
53							4	5.0	86.0
54							4	5.0	91.0
55									
56							1	1.2	92.2
57							2	1.2	93.4
58									
59							1	1.2	94.6
60							4	5.0	99.6
Mean	29.4			34.4			44.2		
S. D.	2.56			2.1			7.00		
M	.37			.39			.79		

The validity of the objective Rorschach is clearly demonstrated by the following critical ratios: normals versus neurotics 9.3; normals versus psychotics 17.0 and psychotics versus neurotics 11.1. There is overlap between the distributions obtained for the three samples. Between normals and neurotics there is a 48.7* percent overlap, between neurotics and psychotics there is a 32.7 percent overlap, and between normals and psychotics there is an overlapping in 2.3** percent of cases tested.

In differentiating between normals and neurotics the establishment of a cut-off point at a score of 33 (median score of overlap range) results in misclassification of 2 percent of the normals and 40 percent of neurotics. If 37 (median score of overlap range) be taken as a cut-off point between psychotics and neurotics, some 3.3 percent of neurotics and 20 percent of psychotics are misclassified. As one would anticipate, the differentiation of normals from psychotics is more clear cut. By taking 34 (median score of overlap range) we find only 2 percent of normals and 2.5 percent of psychotics are wrongly categorized. In this connection, it is anticipated that an item analysis, in which we are presently engaged, will lead to the rejection of response items which are diagnostically equivocal and will thus reduce the degree of overlap which now exists. In any case, in its present form, the test shows overlap only in those cases in which other tests and clinical judgments are equally indefinite, whereas, unlike these, it has provided a rapid, easy and effective method differentiating between the vast majority of cases listed.

Of extreme importance, too, are the upper limit scores of normals and of neurotics. The first of these, a score of 33, indicates the point above which must be regarded as certainly non-normal. The second score 37, the point above which scores must be regarded as almost certainly psychotic.

Summary

This preliminary study describes the application of an objective technique for the administration and scoring of the Rorschach test. This modified technique involves the presentation of the Rorschach cards with instructions to pick two responses which the subject feels best describe the card for him. The results of the preliminary investigation show that the objective Rorschach is a highly valid and reliable method for differentiating diagnostic groups and that this test correlates highly with other diagnostic instruments and subjective clinical assessment. The results also indicate that the test is likely to be an extremely useful instrument in the evaluation of prognosis and in the assessment of therapeutic methods and progress.

Bibliography

1. Baker, L. M. and Creager, J. A. Rating Scale technique applied to Rorschach responses. *J. Clin. Psychol.*, 1954, 10, 373-375.
2. Eysenck, H. J. A comparative study of four screening tests for neurotics. *Psychol. Bull.*, 1945, 42, 639-622.
3. Harrower-Erickson, M. R. and Steiner, M. E. Large Scale Rorschach Techniques. Springfield, Ill.: Chas. C. Thomas, 1945.
4. Munroe, R. L. The Inspection Technique; a method of rapid evaluation of the Rorschach protocol. *Rorschach Res. Exch.*, 1944, 8, 46-70.
5. Steiner, M. E. The Use of the Rorschach Method in Industry. *Rorschach Res. Exch.*, 1947, 11, 46-52.
6. Eysenck, H. J. Dimensions of Personality. London: Routledge and Kegan Paul Ltd., 1947.

---

*As shown in Table 2 there is one abnormally high score of 36 obtained in the normal group. An investigation of this testee revealed her to be in an anxiety state on the day of the test. It is felt that her anxiety affected her performance on the test and led to a higher score. By omitting her score, the percent decreases to 32.5 percent.

**Omission of score 36 decreases the percent to zero.

HOW IMPORTANT IS PSYCHOTHERAPY TO THE HOSPITALIZED PSYCHIATRIC PATIENT[1]

BERNARD A. STOTSKY

Veterans Administration Hospital, Brockton, Massachusetts

PROBLEM

The process of psychotherapy consists of a series of verbal communications between a therapist and a patient. This involves not only the mutual comprehension and manipulation of concepts and symbols but also the ability to utilize words as a medium for expressing emotions. Significant differences have been found between middle class and lower class subjects in preferences for verbal and motoric orientations, with middle class subjects being more verbally and conceptually oriented, while lower class subjects are more motorically inclined.[2]

These findings suggest that psychotherapy should be more suitable for middle rather than lower class patients. Colby[1] and Redlich, Hollingshead, and Bellis[4] note that persons from the lower socioeconomic groups rarely come to psychotherapy spontaneously or voluntarily, communicate poorly with therapists, usually regard their difficulties as purely somatic in nature, and resist any interpretations relating their symptoms to emotional problems. These resistances raise the question of the importance of psychotherapy as a form of treatment from the point of view of the patient. No attempt has been made to assess the attitudes of patients in this respect. This paper is a report of such a survey of the attitudes of a representative group of psychiatric patients at a VA neuropsychiatric hospital. The threefold purpose of this investigation was (a) to determine the degree to which a predominantly lower class group of patients preferred psychotherapy to other therapeutic activities in a mental hospital, (b) to investigate, within the limits of this sample, the relationship between preference for psychotherapy and such variables as occupation, social class, experience with psychotherapy, and degree of chronicity of illness, and (c) to assess patient attitudes concerning the importance and helpfulness of psychotherapy as treatment. The findings of previous studies suggested that the patients sampled here would show less preference for psychotherapy than for other, more motorically oriented activities, and that they would feel that it was less significant than other activities in bringing about improvement in their conditions.

METHOD

Sixty hospitalized NP male war veterans and 30 normal male job applicants served as subjects. Their ages ranged from 18 to 64 with education from fourth grade to three years of graduate work. Seventy-two of the subjects had held as their main jobs some type of manual work, while 18 had held clerical or professional positions. None of the normals had ever received any psychiatric treatment; all scored within normal limits on the Manson Personality Evaluation, a psychiatric screening test. Over 80 percent of the patients were diagnosed as schizophrenics, the remainder as psychoneurotics, behavior disorders, and organics. The normals and patients were equivalent with respect to age, education, and occupation.

Activities Preference Scale (APS). To measure relative preferences for different rehabilitation and treatment activities, a 66 item forced choice preference inventory was devised by the author, consisting of 12 statements referring to major prescribed activities for patients at the hospital. Each statement was paired once with every other one, making altogether 66 pairs. Each patient was instructed to select from each pair the activity he preferred. Thus, a statement would appear eleven times,

[1]The author gratefully acknowledges the assistance of Messrs. Zilaitis, Vardack, Page, Dr. Coburn, and Dr. Conners in collecting the data.

every time paired with a different one. A patient's score was the number of times he selected each of the 12 statements. The statements dealt with:

Working on the grounds outdoors
Sports
Swimming
Work assignment in hospital industries
Woodworking
Working with machines on metal
Doing careful work with tools like in printing and electrical work
Occupational therapy (leatherwork, pottery)
Talking over your problems with the doctor
Operating office machines
Doing bookkeeping in an office
Typing and filing

The first three items refer to outdoor or recreational work, the fourth to regular work assignments, the next four to manual arts activities, the ninth to psychotherapy, and the last three to clerical activities. For each of the twelve a patient's score was the number of times he expressed a preference for the particular activity. The item, "Talking over your problems with the doctor," though more general than a specific reference to psychotherapy (a word not often used by these patients) was selected as the statement to use as a measure of degree of preference for psychotherapy after pretest interviewing showed it to be comprehended by the patients in that light.

The scale was administered individually. Reliability was determined by computing for 20 patients' test-retest correlations for the scores on each of the 12 statements for every subject. The average correlation for the 12 statements was obtained for every patient. Test-retest r's for the 20 patients ranged from .40 to 1.0, with an average of .70.

Results

To determine whether any overall differences existed with respect to preferences for activities, analyses of variance were computed for both the normals and the patients. The F for normals was 37.2, for patients 4.0, both significant at the .001 level. For normals, the greatest preferences were for sports and swimming, with differences between sports and ten of the remaining eleven categories significant at the five percent level. The psychotherapy, bookkeeping, and typing statements were heavily underchosen. For patients, the grounds and hospital industry items were most frequently chosen with psychotherapy and bookkeeping least often selected. For the normals, 47 of the 66 differences were significant; for psychotics, 21.

Comparison of normals and patients revealed significant differences for sports, swimming, woodworking, psychotherapy, office machines, typing and filing, and

Table 1. Activity Preference Scores of 30 Normals and 60 Mental Patients

Activity	Normals Mean	Patients Mean	t	p
Hosp. Industry	6.83	6.60	.4	n.s.
Occup. Therapy	5.03	5.30	.4	n.s.
Swimming	8.37	5.82	3.5	.001
Sports	9.37	5.97	5.7	.001
Grounds	6.87	6.72	.3	n.s.
Woodworking	6.83	5.00	3.2	.01
Metals	5.90	5.50	.6	n.s.
Print. and Elect.	6.53	5.48	1.5	n.s.
Psychotherapy	2.87	4.53	2.6	.01
Office Machines	4.03	5.42	3.0	.01
Bookkeeping	1.47	3.97	3.8	.001
Type and File	1.93	4.90	4.3	.001

TABLE 2. ANALYSIS OF VARIANCE OF SCORES FOR NORMALS AND PATIENTS ON APS

Source	Sum Squares	df	Var. Est
Normals			
Between	2043.4	11	185.8
Within	1729.6	348	5.0
Total	3773.0	359	F=37.2, <.001
Patients			
Between	414.1	11	37.7
Within	6700.7	708	9.5
Total	7114.8	719	F=4.0, <.001

bookkeeping with normals obtaining higher scores on the first three and patients on the last four.

Preference for psychotherapy. In no instance did patients show a significantly greater preference for psychotherapy over any other activity. They did prefer work in hospital industries, on the grounds, sports, and swimming to a significantly greater degree than psychotherapy. Normals chose sports, swimming, hospital industries, grounds, occupational therapy, woodworking, metals, print and electrical work significantly more often than psychotherapy.

The eleven items in which psychotherapy was paired with some other activity were analyzed next to determine patient reaction when the choice was limited to either psychotherapy or some other activity. For ten of the activities a majority of patients preferred the alternative to psychotherapy, differences for two, office work with machines and typing and filing, being significant. The probability of obtaining ten out of eleven preferences in one direction is .017. The findings suggest that psychotherapy is a less preferred form of activity for these patients.

In sampling, two questions arose. What effect did the presence of a psychologist have on the patients' answers? Patients might select the psychotherapy alternative to please the doctor or for fear of reprisal for negative answers. To control for this possible bias, 20 scales were administered by hospital aides and occupational therapists. No consistent differences were found between scores for these scales and those administered by the psychologists.

The second question pertained to the wording of the psychotherapy item. Objection might be raised to the rather broad definition of the term. Therefore, the reactions of 38 patients to the terms "psychotherapy" and "counseling" were obtained and compared with reactions to 27 other hospital activities. Only those patients who understood to what the words referred were sampled. In comparison with manual arts, recreational and work activities, psychotherapy and counseling were overwhelmingly rejected, with differences significant at the .001 level. Only in comparison with clerical activities were the differences not significant, though the trend was toward clerical work. Thus, the wording of the question does not appear to have contributed significantly to the low preference for psychotherapy.

Relation of various factors to APS scores. The first comparison was between the 72 manual workers and the 18 clerical and professional workers. None of the differences between these two groups were significant. Of the 18 clerical workers all but six came from lower class families so that they might be expected to show the influence of early training somewhat in their preferences. The six supplemented by 12 other patients with middle class backgrounds were compared with the other 84 for possible differences. Comparisons were made by means of *chi square* using the median of the entire sample as a cutting point. Differences for the "outdoor working on the grounds" and psychotherapy items were significant with lower class patients preferring the former, middle class patients the latter.

Patients receiving psychotherapy were compared with patients not receiving it. No significant differences were found in preferences for any of the 12 activities. A further categorization was made in terms of severity of illness. The patients were divided into those with illnesses of short duration, chronic but stabilized, and acutely disturbed patients. Comparison of the three groups for all activities failed to yield any significant differences.

Patient attitudes concerning importance and helpfulness of psychotherapy. Although the patients had been given an opportunity to express their preferences for psychotherapy relative to other activities, they had not been asked to indicate whether they felt that it had contributed to making them well. To obtain data on this question, 30 patients undergoing individual psychotherapy were asked whether they felt that psychotherapy was the most important part of their treatment. Twenty three answered "No". Sixteen of the 30 felt that psychotherapy had helped. Since simple "yes - no" alternatives to questions are not regarded as giving a true picture of the salience of attitudes[3], an attempt was made to assess the readiness of patients to regard psychotherapy as significant in their treatment. Sixty patients were asked the following questions:

> "Which of the activities or treatments at the hospital have helped you to get better?" "Of these which was most important?"

Only six patients (all of whom were in psychotherapy) spontaneously mentioned psychotherapy as having helped. One regarded it as most important. Twenty-six of the 60 patients were receiving some form of psychotherapy. Forty-one of the patients mentioned rehabilitation and recreational activities as helpful; 10 mentioned drugs.

Discussion

The normals and patients in this predominantly lower class sample were much less favorably disposed to psychotherapy than to physical recreation, manual work, and activities involving work by hand, such as woodworking, metal work, electrical and printing work. The normals, even more than the patients, tended to reject psychotherapy. Further breakdown of the data showed no significant relationship between occupation and activity preferences. Significant differences were found between patients from lower and patients from middle class families for the "grounds" and psychotherapy items. Differences in the severity of illness were not related to activity preferences. In the minds of the great majority of patients psychotherapy did not play a dominant even an important role in their treatment. They regarded recreational and rehabilitation activities as more significant and more helpful in their recovery than either psychotherapy or chemotherapy. The patients preferred to think of their treatment in terms of activities other than psychotherapy. This was true even of the majority of patients receiving psychotherapy. Our findings parallel those of McNeil with respect to the primarily motoric orientation of lower class individuals and those of Redlich, Hollingshead, and Bellis concerning the attitudes of lower class patients toward psychotherapy.

The findings suggest that it may be necessary to develop new techniques for treating lower class patients which make use of the primarily motoric orientations of these patients. Such a form of treatment should be tied in closely with rehabilitation and recreational activities which the patients prefer. Two studies[5, 6] conducted at this hospital with chronic regressed and disturbed schizophrenics have demonstrated the value of integrating psychotherapy and counseling with work-oriented, rehabilitation programs. As the techniques of psychotherapy are modified to make them more meaningful to the less verbal patients, it may become possible to treat large groups of patients now regarded as relatively inaccessible to psychological treatment.

Summary

A predominantly lower class sample of patients at a VA neuropsychiatric hospital was surveyed to assess attitudes towards psychotherapy. Two types of data were secured: (a) an activities preference questionnaire designed to measure relative preferences for 12 rehabilitation and therapeutic activities; (b) responses to structured and unstructured questions relating to activities and modes of treatment regarded as contributing to improvement in their illnesses. As expected, psychotherapy was regarded as less important than manually oriented activities. It was mentioned spontaneously by only ten percent of the patients as contributing to their getting well. Patients with middle class backgrounds showed a greater preference for psychotherapy than patients with lower class backgrounds. Neither occupation nor severity of illness showed any relationship to activity preferences. The implications of the findings were discussed.

References

1. Colby, K. M. *A primer for psychotherapists.* New York: Ronald Press, 1951.
2. McNeil, E. B. *Conceptual and motoric expressiveness in two social classes.* Unpublished doctor's dissertation, Univ. of Michigan, 1953.
3. Newcomb, T. M. *Social Psychology.* New York: Dryden Press, 1950.
4. Redlich, F. C., Hollingshead, A. B., and Bellis, E. Class differences in attitude toward psychiatry. *Amer. J. Orthopsychiat.*, 1955, 25, 60-70.
5. Stotsky, B. A. Predicting success on the member-employee rehabilitation program. *J. consult. Psychol.*, in press.
6. Stotsky, B. A., Daston, P. G., and Vardack, C. N. An experimental evaluation of counseling of regressed and disturbed schizophrenics. *J. counsel. Psychol.*, in press.

SETECTION OF ABBREVIATED TAT SETS[1]

RICHARD H. DANA

St. Louis State Hospital

Problem

Short forms of the TAT are routinely and unquestioningly employed in clinical psychological practice. The basis for selection of appropriate cards has always been "clinical utility", *i.e.*, those cards which will supposedly "pull" or elicit relevant information in particular areas of personality. This results in reliance upon informed clinical bias in card selection. Problems of card selection also occur in clinical research projects where demands of time and economy in test battery construction often dictate choice of a few TAT cards.

Evaluation can proceed by recognition of two related questions: (a) Should short forms of longer tests be used at all? Are they reliable, valid, and clinically important? (b) Is card pull *per se* a sufficient, empirical criterion for selection? These questions will be considered both historically and experimentally.

Current Status of Short Form Tests

Since the composition of short form TAT sets is part of the larger issue of short form tests in general, some of the literature on equivalence of information from short forms and full scale tests, both psychometric and projective, will be considered. This discussion will be limited to shortened forms of standard tests and will not include brief tests.

[1]This investigation was supported in part by a research grant from the National Institute of Mental Health of the National Institutes of Health, Public Health Service. Grateful acknowledgment is made to Dr. Lewis Bernstein and to the psychology and research staff of the St. Louis State Hospital for cooperation.

Three separate studies on the Wechsler-Bellevue attest to the valid relationships between short forms and full scale which may be empirically established[19, 20, 31]. Classification tests, selected from batteries, and their length halved are of equivalent reliability and validity with full length tests[6]. Shortened MMPI forms, group and individual, appear of more questionable validity[27, 28]. Thus, when objective test scores are available, short forms of some longer tests may be empirically feasible and of demonstrably similar usefulness.

Short forms of projective techniques have not received similar research attention. Single Szondi administrations are unreliable[7]. Dorken's single ink blot test prepared by the subject[14, 15] and Zulliger's three card version of the Rorschach are most properly described as brief tests[3, 24, 32]. Valentine[38], using Beck's 47 published Rorschach protocols[4], abstracted the scores for a four plate form (II, IV, VI, X) and a five plate form (II, IV, VI, VIII, X) and correlated these with scores from the full set of ten plates. Correlations between abridged scores and total scores for 12 main scoring variables ranged from .48 to .93 with four plates and from .83 to .98 with five plates. Variability of scoring was not appreciably increased using the abridged set. The abridged set was then administered followed by the other plates to a new sample of 60 subjects. A mean correlation of .93 was obtained between abridged scores and total scores.

Although Valentine's study has been largely ignored, it is a matter of evidence that a complete Rorschach evaluation requires "a little more than four hours"[33] and is considered insufficient as the sole diagnostic instrument[36]. The effect of position or order of presentation has been studied for the Rorschach[30] as well as for personality inventories[18] and the necessity for restandardization of any altered test is recognized[37, 43].

The basic question, "How long is long enough?", has been greeted with scandalized sensibilities and refusal to admit of small samples used for clinical research[35], by the considered statement that there is no currently satisfactory basis for an answer[29], and by acceptance of "every single response and every part of a response as significant and representative of the subject[34, p. 40]".

It seems evident from even these meager research results that small samples of behavior, i.e., short forms of projective tests, can be usefully employed. Statistical inroads are continually being made on problems of optimal and standard test length [21, 22, 23]. Whenever short forms are used, the assumption must be made that short form and full length test are roughly equivalent behavior samples at least with respect to clinically meaningful objective scoring. This assumption should be empirically tested for any given scoring system.

Card Pull and Card Selection

"Card Pull" can be defined as the amount of relevant personality data elicited. The literature on short form TAT sets has been reviewed by Watson[39] who cites six different sets composed on a combined *a priori* and empirical basis of either 9 or 12 cards and developed by Eiduson and Klopfer, Garfield and Eron, Holt, Levinson, Rosenzweig, and Watson. When these selections are summarized, complete agreement is found for male series cards 3BM, 4, 13MF, and 18BM, and all but one agreed upon cards 6BM, 7BM, and 12M. Complete agreement is found for female series cards 4 and 13MF; all but one included cards 1, 2, 15, and 18GF. One investigator[8, 11] has employed short forms chosen to compose the fundamental interpersonal life situations.

Exploration of the TAT card selection criterion, amount of relevant personality data elicited, has involved experimental evaluation of the subtle interplay of degree of stimulus ambiguity and amount of personality data elicited[5, 25, 26, 40, 41, 42]. Impetus to research was provided by development of the Transcendence Index, the mean N of comments going beyond pure card description[40]. As ambiguity increases so does the amount of projection or fantasy measured by this index. Those cards with the highest Transcendence Indices (6 BM, 4, 7BM, 7GF, 2) were also the least ambiguous stimuli. However, when psychological ambiguity was employed, i.e., con-

sensus on estimated N of possible interpretations from each card, those cards of medium ambiguity (4, 7BM, 8BM, 12M, 13MF) yielded the greatest amount of personality information in terms of clinical judgment[26]. In spite of differences in definition and measurement of both stimulus ambiguity and personality data, there are marked similarities in cards selected. In general, the assumption that Transcendence Index and amount of personality data are related has been supported. However, relationships between physical ambiguity and transcendence are still not clearly delineated. In Kenny's words, "it may be hypothesized that a decreasing number of cue constellations available for selective reaction produces response tendencies which do not contain much in the way of individual projections[25, p. 347]." If this be valid, then short form TAT sets could readily be assembled from among those cards presenting the most numerous stimulus cues.

Stimulus Cues and Personality Data

The problem of establishing relationships between TAT stimulus cues and individual projections can be experimentally investigated. The N of physical stimulus cues was determined by having ten trained clinicians list the cues from each of the TAT cards and the mean N of these cues was obtained. The N of individual projections, the personality data variable, was defined as the N of themes produced to a given card. This data was available from Eron's studies in terms of the total N of themes, the N of different themes, and the N of unusual formal characteristics contributed by two or more women *Ss*[16], and the total N of themes for male *Ss*[17].

Rank-difference correlations indicated no significant relationships between stimulus cues and themes or unusual formal characteristics (Table 1). There was,

Table 1. Rank-Difference Correlations Between Individual Projections and Stimulus-Cues for 19 TAT Cards

Factors		Male	Female		
		N themes	N themes	N diff. themes	unus. char.
N of S-cues	Male	−.06			
	Female		.22	−.09	−.12

however, a significant relationship evinced between the N of stimulus cues present in cards of the male and female series (.70), although this did not occur for the N of themes in male and female cards (.32). These results suggest several alternative considerations: (a) the use of themes and unusual formal characteristics may not be an entirely appropriate measure of the amount of personality data elicited; (b) the mere tally of physical stimulus cues may not yield an adequate or defensible measure of TAT stimulus value; (c) the attempt to establish quantitative relationships between specific card properties and personality data although experimentally feasible may be clinically naive. In addition, these negative results emphasize dangers involved in the oversimplification implied by any one-to-one relationships in the area of personality study. Short form TAT sets clearly cannot be developed solely upon the basis of such criteria.

Objective Scoring and Short Form TAT Sets

Objective TAT scoring that is economical, reliable, and valid for gross screening purposes has only recently been introduced[8]. This TAT scoring was developed from a consistent approach to personality theory[9] and from a rationale for application to various projective techniques[10]. Currently, this scoring system is providing descriptive personality data in addition to degree of psychopathology[13], and has been validated for use with female *Ss*[11], and cross validated on male *Ss*[12].

One scoring category, Perceptual Organization (PO), which alone provides considerable information with regard to psychopathology and personality description,

is being used here in comparison of short form sets of TAT cards. PO is defined as the degree to which standard test directions are followed in the *S*'s TAT story. The PO components are card description, present behavior, past events, future events, feeling, thought, and outcome. One point is scored for each part used in the story.

The TAT protocols (20 cards of F, GF series) of 67 normal women were scored for PO. Total scores and mean scores were obtained. The PO scores for each of the 20 cards were correlated with the total score for all cards (range .60 to .78), and a mean correlation determined by transformation to Z scores (.70). The mean scores of short form sets selected by various investigators for women on *a priori*, empirical, or experimental grounds as having the greatest "card pull" were correlated with the mean scores for the total set. The mean scores of a short form set selected at random were also correlated with the mean scores for the total set (Table 2). These cor-

TABLE 2. CORRELATIONS OF PO SCORES ON PROPOSED SHORT FORM TAT SETS WITH TOTAL SCORES FOR ALL 20 CARDS.

Set	Cards used	*r*
Kenny	4, 13MF	.82
Weisskopf	2, 4, 7GF	.85
Dana	2, 3GF, 4, 6GF, 7GF	.91
Watson et al.	1, 2, 4, 13MF, 15, 18GF	.93
Random	1, 3GF, 12F, 14, 17GF	.95

relations appear to be dependent largely on the N of cards employed in the short form, with the possible exception of the random set. When TAT stories are scored for Perceptual Organization it does not seem to make any real difference which set of cards is used. It is important how many cards are included. A female sample and cards of the F, GF series limited the number of available short form sets which could be evaluated. However, notwithstanding this fact, the relevance of this study rests on the *validity* of the objective scoring category employed. It must be emphasized that validity for objective scores encompasses not only prediction of clinical diagnosis (degree of psychopathology) but personality description as well. This exemplifies a methodology which should be applied with various other objective scoring categories to determine whether or not meaningful generalizations can be made from these results.

SUMMARY AND CONCLUSIONS

A review of pertinent literature and two brief experimental demonstrations provided some information on the selection and evaluation of short form TAT sets.

(1) Short form TAT sets, or short forms of other projective techniques, can be evaluated for *any* available objective scoring categories by statistical comparison with the total set of cards.

(2) TAT sets of as few as five cards, when scored for Perceptual Organization, provide data which are roughly equivalent to that from the entire series.

(3) The inclusion of particular cards appears of less importance than the actual number of cards selected.

(4) Choice of specific cards in terms of a quantitative criterion of stimulus cues alone is not adequate for development of short form TAT sets.

(5) Although "card pull" and card selection are undoubtedly related, more precise and consensually valid definition of what constitutes "relevant personality data" must inevitably precede measurement of this value for each TAT card and empirical evaluations of cards thereby selected.

REFERENCES

1. ALLISON, R. B., JR. Battery validity as a function of the time devoted to each test. *Educ. psychol. measmt.*, 1953, 13, 288-296.
2. AYERS, J. D. A simple technique for estimating time limits. *Canad. J. Psychol.*, 1953, 7, 81-85.

3. Bacci, V. Prime considerazione sull'applicizione in Italia del reattivo di Zulliger. *Riv. Psicol.*, 1949, 45, 138-142.
4. Beck, S. J. *Rorschach's Test. II. A variety of personality pictures.* New York: Grune and Stratton, 1949.
5. Bijou, S. W. and Kenny, D. T. The ambiguity values of TAT cards. *J. consult. Psychol.*, 1951, 15, 203-209.
6. Brokaw, L. D. Comparative validities of "short" versus "long" tests. *J. appl. Psychol.*, 1951, 35, 325-330.
7. Cole, D. The reliability of a single Szondi profile. *J. clin. Psychol.*, 1951, 7, 383-384.
8. Dana, R. H. Clinical diagnosis and objective TAT scoring. *J. abnorm. soc. Psychol.*, 1955, 50, 19-25.
9. Dana, R. H. Personality Orientation: an organizational focus for current research. *J. Psychol.*, 1954, 37, 139-150.
10. Dana, R. H. The objectification of projective techniques: rationale. *Psychol. Rep.*, 1955, 1, 93-102.
11. Dana, R. H. An application of objective TAT scoring. *J. proj. Tech.*, in press.
12. Dana, R. H. Cross validation of objective TAT scoring. *J. consult. Psychol.*, in press.
13. Dana, R. H. The Tatogram: a case study. (To be published).
14. Dorken, H., Jr. The ink blot test; clinical application of a brief projective test. *Rev. Psychol.* 1952, 1, 443-468.
15. Dorken, H., Jr. The ink blot test; clinical application of a brief projective technique. *Rorschachiana*, 1952, 1, 196-221.
16. Eron, L. D. Responses of women to the TAT. *J. consult. Psychol.*, 1953, 17, 269-282.
17. Eron, L. D. A normative study of the TAT. *Psychol. Monogr.*, 1950, 64, (9), No. 315.
18. Gordon, L. V. The effect of position on the preference value of personality items. *Educ. psychol. measmt.*, 1952, 12, 669-676.
19. Herring, F. H. An evaluation of published short forms of the Wechsler-Bellevue scale. *J. consult. Psychol.*, 1952, 16, 119-123.
20. Hilden, A., Taylor, J. W. and Dubois, P. H. Empirical evaluation of short W-B scales. *J. clin. Psychol.*, 1952, 8, 323-331.
21. Horst, P. Determination of optimal test length to maximize the multiple correlation. *Psychometrika*, 1959, 14, 79-88.
22. Horst, P. Optimal test length for maximum validity. *Psychometrika*, 1951, 16, 189-202.
23. Horst, P. A note on optimal test length. *Psychometrika*, 1950, 15, 407-408.
24. Jaur, J. M. Etude experimentale sur le test de Zulliger. *Arch. Psicol. Neurol. Psychiat.*, 1953, 14, 159-166.
25. Kenny, D. T. Transcendence indices, extent of personality factors in fantasy responses, and the ambiguity of TAT cards. *J. consult. Psychol.*, 1954, 18, 345-348.
26. Kenny, D. T. and Bijou, S. W. Ambiguity of pictures and extent of personality factors in fantasy responses. *J. consult. Psychol.*, 1953, 17, 283-288.
27. MacDonald, G. L. Effect of test-retest interval and item arrangement on the shortened forms of the MMPI. *J. clin. Psychol.*, 1952, 8, 408-410.
28. MacDonald, G. L. A study of the shortened group and individual forms of the MMPI. *J. clin. Psychol.*, 1952, 8, 309-311.
29. MacFarlane, Jean W., and Tuddenham, R. D. Problems in the validation of projective techniques. In *An introduction to projective techniques.* H. Anderson & Gladys Anderson (Eds.) New York: Prentice-Hall, 1952.
30. Maradie, L. J. Productivity on the Rorschach as a function of order of presentation. *J. consult. Psychol.*, 1953, 17, 32-35.
31. McNemar, Q. On abbreviated Wechsler-Bellevue scales. *J. consult. Psychol.*, 1950, 14, 79-81.
32. Morali-Daninos, A., Canivet, N., and Thierry-Mieg, B. Le test "Z" en psychotechniques. *Travail hum.*, 1952, 15, 277-291.
33. Odom, C. L. A study of the time required to do a Rorschach examination. *J. proj. tech.*, 1950, 14, 464-468.
34. Rapaport, D. *Diagnostic psychological testing. Vol. I.* Chicago: Year Book, 1945.
35. Rotter, J. B. Clinical methods: psychodiagnostics. In *Annual Review of Psychology.* C. P. Stone (Ed.) Stanford: Stanford Univ. Press, 1953.
36. Sacks, J. M., & Lewin, H. S. Limitations of the Rorschach as sole diagnostic instrument. *J. consult. Psychol.*, 1950, 14, 479-481.
37. Technical recommendations for psychological tests and diagnostic techniques. *Psychol. Bull. Supp.*, 1954, 51, No. 2, Part 2.
38. Valentine, M. Experimental abridgement of the Rorschach test. *J. clin. exp. Psychopath.*, 1951, 12, 157-169.
39. Watson, R. I. *The clinical method in psychology.* New York: Harpers, 1951.
40. Weisskopf, Edith A. A transcendence index as a proposed measure in the TAT. *J. Psychol.*, 1950, 29, 379-390.
41. Weisskopf, Edith A. An experimental study of the effect of brightness and ambiguity on projection in the Thematic Apperception Test. *J. Psychol.*, 1950, 29, 407-416.
42. Weisskopf-Joelson, Edith A., and Lynn, D. B. The effect of variations in ambiguity on projection in the Children's Apperception Test. *J. consult. Psychol.*, 1953, 17, 67-70.
43. Wilson, J. W., and Carpenter, E. K. The need for restandardization of altered tests. *Amer. Psychologist*, 1948, 3, 172-177.

THE RELATIONSHIP BETWEEN CERTAIN PARENTS' ATTITUDES TOWARD CHILD REARING AND THE DIRECTION OF AGGRESSION OF THEIR YOUNG ADULT OFFSPRING

FRIEDA L. BORNSTON AND JAMES C. COLEMAN

Los Angeles City College, Evening Division *University of California at Los Angeles*

Problem

The crucial importance of parent attitudes in shaping the personality development of the child has been repeatedly emphasized. Although this emphasis has some justification in experimental studies, it is largely based upon qualitative cues derived from clinical observation and experience. The present study is an attempt to extend experimental information in this area. Specifically, it is a study of the relationship between certain attitudes of the mother and the direction of aggression elicited in her offspring by mildly frustrating situations.

Method

For purposes of the present study, two test instruments were selected, the adult form of the Rosenzweig Picture-Frustration Study (P-F) and the University of Southern California Parent Attitude Survey (P-A-S). Also a personal data sheet was used for obtaining certain additional information about the subjects.

This test battery was administered to 152 college students. Special directions given with the P-A-S requested that each student answer the questions as they believed their mothers would. A second copy of the P-A-S was given to the students with a request that they have their mothers complete it and return it to the examiner via the mails in the stamped, addressed envelope provided for this purpose. Complete anonymity was assured, both to the students and their parents, by a numbering system which still permitted the pairing of mothers and children for statistical handling of the data. This procedure yielded 56 complete sets of tests for scoring. The P-A-S was scored according to the weighted scoring devised by Shoben[2], and the P-F according to a new method devised for purposes of this study.[1]

Of the 56 students whose responses figure in the final results of this study, 24 were men and 32 were women. The mean age was approximately 20 years with a range from 17 to 40. The great majority were native-born, white, Protestant Americans. The ages of the 56 mothers who participated in this study ranged from 36 to 70 years with a mean of approximately 42 years. The mean educational level was the eleventh grade with a range from the sixth grade to college graduation. The average family income as reported by students and parents approximated $5000 per year. Here too, the great majority of mothers were white, Protestant, native-born Americans.

The tests yielded four sets of scores: (a) the actual maternal scores on the P-A-S, (b) the children's estimates of maternal scores on the P-A-S, (c) the difference scores between the actual maternal P-A-S scores and the children's estimates of them which were taken as a measure of the child's perceptual distortion of maternal attitudes, and (d) the children's scores on the P-F.

Each of the sets of scores on the P-A-S consisted of scores on a total scale derived by summing the scores of four sub-scales: (a) dominance, (b) possessiveness, (c) ignoring, and (d) unclassified. According to Shoben[2], the higher the score on the dominance scale, the greater the tendency to consider a child as one who should conform completely to parental wishes under penalty of severe punishment. The higher the score on the ignoring scale, the greater the tendency to disregard the child as an individual member of the family. The higher the score on the possessive scale, the greater the tendency to "baby" the child, to emphasize affectional bonds between

[1]Discussed and evaluated in an unpublished study.

parent and child and to value the child's dependence on the parent. Shoben does not delineate the meaning of the total or the unclassified scales. The former, as stated, represents a simple sum of the sub-scales. The latter includes items which attempt to tap attitudes toward a variety of factors including sexual activity, religious and/or moral behavior, and money matters.

The P-F provides three scores for each individual for the three possible directions of aggression: (a) extrapunitive (*E*), (b) intropunitive (*I*), and (c) impunitive (*M*). Both the extra- and intropunitively directed aggression are seen by Rosenzweig as stemming from a basic aggressive drive. The impunitive score, however, is seen as arising from erotic drives. An additional score was, therefore, computed by combining the E and I scores as a measure of what might be called "hostile" aggression.

There were, thus, four scoring categories on the P-F. Each of these was correlated with each of the scales on the three sets of Parent Attitude Surveys obtained.

Results

Results of the study are summarized in Table 1. It may be seen that, of the fifty correlation coefficients computed, nine were found to be significant at greater

Table 1. Pearson Correlation Coefficients Computed Between Maternal Scores, Children's Estimates and Difference Scores on Parent-Attitude-Survey and Children's Scores on P-F (N=56)

Children's Scores on P-F	Factors	Total	Sub-Scales on P-A-S Dominant	Possessive	Ignoring	Unclassified
E	Maternal Scores on P-A-S	.240	.263*	.094	.186	.297*
	Children's Estimate	−.002	−.007	−.006	−.009	−.031
	Difference Scores	.008	.048	−.008	.178	.140
I	Maternal Scores	.006	.006	.121	.112	−.022
	Children's Estimate	−.002	−.004	−.015	−.066	−.0009
	Difference Scores	−.003	.024	.021	−.168	−.064
M	Maternal Scores	−.280*	−.292*	−.215	−.251	−.273*
	Children's Estimate	.003	.004	.010	.017	.015
	Difference Scores	.155	−.173	−.070	−.099	−.146
E I	Maternal Scores	.273*	.291*	.183	.325*	.319*
	Children's Estimate		(Not	Computed)		
	Difference Scores					

*Significance > .05 but < .01 level of confidence.

than the five percent but less than the one percent level of confidence. All of the significant correlations occurred between the children's P-F scores and the actual maternal scores on the P-A-S. These were: 1. A positive correlation between the child's extrapunitiveness and the mother's dominance and unclassified scores. 2. A negative correlation between the child's impunitiveness and the mother's total, dominance and unclassified scores. 3. A positive correlation between the child's hostile aggression (E plus I) score and the mother's total, dominance, ignoring, and unclassified scores.

Discussion

There are three main findings which emerge from the present study. The first is that the child's interpretation of his parents' attitudes or the extent that this interpretation is a distortion of actual parental attitudes does not appear to be significantly related to the direction of h s aggressive responses under stress. This is contrary to the assumptions made by many clinical psychologists that it is the child's view of parental attitudes rather than the attitudes per se which are important determiners of personality characteristics.

The second finding is that a predictive index of a child's aggressive reactions to stress may be obtained from the parent's own statement of attitudes. This statement holds for two of the dimensions of parental attitudes measured in the present study, the dominant and ignoring attitudes. The fact that these are significantly related to the child's aggressive reactions to frustration supports the interpretation that dominant and/or ignoring mothers arouse hostility in their children but do not enable them to direct this hostility toward them so that the children direct it either inwardly upon themselves or outwardly upon other objects or persons. The latter would appear to be the favored pattern. The third finding indicates that the Unclassified Scale of the P-A-S should be subjected to further study so that its meaning can be more clearly delineated. As it stands, the significant correlations obtained between it and the various reactions to stress cannot be interpreted.

In evaluating the present findings, it should not be overlooked that the present study rests on certain assumptions which underlie its experimental procedures. These assumptions may themselves be open to question and should, perhaps, be made explicit. It was assumed that selection factors in sampling would not invalidate findings that attempt to establish relationships rather than norms. It was further assumed that there is an essential continuity in parental attitudes and that, consequently, the assessment of a parent's attitudes at any given point in time will be an adequate measure of that parent's attitudes over an extended period of time. Although this assumption is often made in theories of personality and has some clinical and observational justification, it is probable that some parents change their views on parent-child relationships over a period of years. Until this point has received experimental clarification, the extent of error introduced by our present assumption of stability of attitudes cannot be accurately assessed.

The study further assumes that individuals are consistent in the direction of their aggressive responses to specific frustrating situations. Although there is, again, clinical evidence to suggest that individuals have characteristic modes of reaction to frustration, there is still the question of the extent to which these are a function of the type of stress situation selected, i.e. possible differences in the modes of reaction to frustrations induced by "superiors" or "inferiors" is not taken into account.

The crucial assumption of the study, however, lies in the assumed validity of its measuring instruments. This, of course, is a characteristic reality factor in much clinical research, but it must be observed that while the mode of validation of the P-A-S is statistically acceptable, it does not insure that the parent who obtains a high score on the dominance scale, for example, will actually display more domineering behavior toward her child than does one who receives a low score on this scale. Similarly there is no conclusive evidence that the individual who responds in a given direction to a P-F item will react in the same fashion in an actual social situation such as its depicts.

Despite these limitations, the present findings are of interest in that they afford experimental support for some and non-support for other prevailing clinical opinions on the nature of the links between development of personality characteristics in the child and his parents' attitudes toward him.

Summary

The present investigation was concerned with the relationship between certain attitudes of the mother as a parent and the direction of aggression elicited in her college offspring by mildly frustrating situations. The major findings were as follows:

1. The greater the tendency of the mother to be domineering, the greater the tendency of the child to react to frustration by hostile aggression directed either outwardly on to the environment or inwardly toward the self.

2. The greater the tendency of the mother to be domineering, the less the tendency for the child to gloss over frustration or evade the expression of aggression.

3. The greater the tendency for the mother to be ignoring, the greater the tendency for the child to respond with hostile aggression to frustration directed either outwardly on to the environment or inwardly toward the self.

4. No significant relationship was found to exist between the child's interpretation of his mother's attitudes and the direction of his aggressive responses to frustrating situations.

5. A mother's own estimates of her attitudes provides a better predictive index for her child's aggressive reactions to frustration than does the child's interpretation of these attitudes.

In general, a statistically significant though low degree of relationship was found between measures of maternal attitude and measures of the direction of aggression elicited in their offspring by frustrating situations resembling those of everyday stress.

Bibliography

1. Rosenzweig, Saul, Fleming, E. E., and Clarke, H. J. Revised scoring manual for the Rosenzweig Picture-Frustration Study. *J. Psychol.* 1947, 24, 165-208.
2. Shoben, Edward J., Jr. The assessment of parental attitudes in relation to child adjustment. *Genet. Psychol. Monogr.*, 1949, 39, 101-148.

TWO TYPES OF PERCEPTUAL EXPERIENCE

NOEL JENKIN[1]

University of New Brunswick

Problem

Clinical and experimental psychologists are both familiar with a broad category of perceptual or interpretative response which has as its main feature a high degree of discrepancy with the objective nature of the stimulus material. Depending on the viewpoint or area of specialization of the experimenter, responses of this type are described as "pre-recognition hypotheses", "non-veridical responses", "perceptual distortions" or "projective responses". Occasional use has been made of such data for special purposes (e.g.[4]) but in general they have been little studied outside the conte t of "projective" personality testing. Rubinstein[6] has employed data of this kind as the dependent variable in an experiment which supported the hypothesis that "perceptual distortion" increases as "personality integration" decreases. The present author has reported evidence[2, 3] which supports this position, but only insofar as a particular type of non-veridical response is concerned, i.e. that termed "Rationalization", characterized by hesitancy and tentativeness in verbal expression. An alternative type of non-veridical response was described as "Projection" involving a positive, definite, unequivocal report, and responses of this kind predominated in the protocols of secure and well-adjusted people.

A question left unanswered by this study, however, was the important issue of whether the types of response distinguished reflect differences in the nature of the perceptual experience itself, or merely represent differences in the manner of reporting the experience. Inspection of the protocols from the experiment cited, together with introspective accounts by some of the subjects, suggested that these differences in the nature of the verbal report, did in fact correspond with differences in the nature of the perceptual experience. If this should be the case, the distinction proposed would have considerable theoretical importance, for both perceptual theory and per-

[1]The data of the investigation reported here were obtained when the author was at Victoria University College, New Zealand. He is indebted to the Laboratory of Social Relations, Harvard University, for facilities to process the data, and to several members of its staff for valuable advice.

sonality theory, and might lead to a new variable of potential value for clinical diagnosis. The following experiment was therefore designed to answer this problem.

METHOD

By means of the low-intensity, veiling-glare technique described in previous reports[2, 3], ambiguous picture material was shown to 89 students of a class in elementary psychology. The experiment was introduced as part of the practical work of the course, but the best cooperation of the students was requested since the data were to be used for research purposes. Answer booklets were handed out, each of which was covered by a blank page. Subjects were asked not to turn the page until told to do so. Instructions for the first part of the experiment were as follows: "When I say 'ready', please watch the screen. Something will then appear on it for several seconds, and afterwards I want you to write down on the blank page covering your booklet, everything you saw and what you think it might be. If the significance of what you saw isn't clear to you I would like you to write down any guesses you make as to what it might be. After your first observation we will repeat the same procedure a number of times. At first, the projection will be very faint, but nevertheless each time I want you to report every impression you have and your belief or guess as to what it is that you see." After questions had been answered, the room was darkened and the first exposure made. Sufficient light was introduced into the room between exposures to allow subjects to write their reports, ten of which were made upon the same slide at progressively increasing levels of illumination. At the first step of illumination, only shadowy masses were visible. At the final step the essential features of the picture were visible, but due probably to fixation of incorrect hypotheses, a considerable number of subjects still failed to yield a correct report.

The second part of the experiment was conducted immediately following the first, without a break. Instructions were: "Now turn over to the next page of your booklet. It consists of a form containing five sections. After each observation which you make this time, I want you to consider carefully which is the most appropriate section to describe your experience. Section I is labelled: 'Nothing at all except illuminated screen.' Make a check mark in this section, in the appropriate column, if this describes your experience after Exposure 1. Section II is labelled: 'Nothing except patches of light and shade, or lines, dots, angles and/or geometrical forms.' If this is the best description of your experience, out of the five alternatives, make your check mark here.

Section III is more elaborate. Please follow the instructions carefully. There are two alternatives here. You can say 'the projection looks like . . . ', writing your impression against the appropriate exposure number, and writing 'A' in the corresponding cell on the right. Or, on the other hand, you can say, 'the projection reminds me of and could possibly represent . . . . ', writing your impression against the appropriate exposure number, and writing 'B' in the corresponding cell on the right. If your impression comes to you vividly, immediately and spontaneously, it will be an 'A' answer. If it comes after some delay, or with some effort on your part, it will be a 'B' answer. Please try to classify your 'A' and 'B' experiences as well as you can.

Section IV is labelled: 'The projection was the same (or essentially the same) as last time, or was clearer but otherwise the same.' Make a check mark here if this best describes your experience, out of the possibilities mentioned on the form. Section V says: 'Although I answered Sections III or IV on the last exposure (in this series), this time I could not say what the projection is like or what it could possibly represent.' Again you should check this section, in the appropriate column, if this describes your experience better than any of the other sections."

After questions had been answered, a series of 10 exposures of a different slide was presented at gradually increasing steps of illumination. Three more series, with a different slide each time, completed Part II of the experiment. The subjects were then asked to fill in a questionnaire which dealt with their satisfaction or dissatisfaction with the record blank as a method of classifying their experiences.

Hypotheses

It was expected that if Projection and Rationalization are distinct kinds of perceptual experience, rather than merely representing differences in the manner of communicating a report, there should be a positive correlation between the number of responses scored as Rationalization by the experimenter in the first part of the experiment and the number so scored, in effect, by the subject himself, in the second part. Similarly, there should be a positive correlation between Projection scores in Parts I and II. Section III A represented Projection responses in Part II, and III B represented Rationalization responses. Such a correlation could also be due to general productiveness on the part of the subject. If, however, there were no significant correlations between Rationalization and Projection either within or between Parts I and II, this evidence would be consistent with the view that these categories represent different kinds of perceptual experience.

Auxiliary support for the same proposition was expected from answers to the questionnaire. If the modes of perceptual experience distinguished by the present hypothesis are in fact "types" rather than extremes of a continuum, the majority of subjects would indicate satisfaction with the qualitative distinction made by the record blank, rather than express a preference for rating their responses on a 5 point scale implying a quantitative distinction. The majority of subjects would also deny that they experienced difficulty in classifying their responses as Type A (Projection) or Type B (Rationalization) and affirm that their responses seemed to fall naturally into the two categories.

Results

The protocols from Part I of the experiment were scored in the following manner. Descriptions of the image on the screen which merely mentioned patches of light and shade or geometrical figures were excluded. The remaining non-veridical responses were then classified, firstly, as "Projections" if the statements were unequivocal, secondly as "Rationalizations" if the statements contained evidence that the subject was guessing, or was doubtful about the objective accuracy of his report. All statements preceded by expressions such as "It could be . . . ", "It might be . . . ", "It suggested . . . ", "It reminds me of . . . ", were scored as Rationalizations.

The responses recorded in Part II of the experiment were then examined, to determine which were non-veridical, i.e. having no correspondence with the objective nature of the stimulus material. Such responses were then tabulated in accordance with the subject's own classification as Type A (Projection) or Type B (Rationalization). Repetitions of previous responses were not counted again in the scoring, either in Part I or in Part II.

The six intercorrelations among the measures on Part I and Part II are shown in Table 1. Projection in Part I correlates with its counterpart in Part II to a highly significant degree. Rationalization in Part I does likewise with its counterpart. All other correlations are below the .05 level of s gnificance. The absence of any significant correlation between unlike pairs of these measures implies that the relationships between like pairs cannot be accounted for in terms of "productiveness." Sup-

Table 1. Correlations Between Scores on Two Forms of Test (N = 89)

	Part I		Part II
	Projection	Rationalization	Projection
Part I Rationalization	−.06	—	
Part II Projection	.44*	−.01	—
Rationalization	.10	.40*	−.14

*P <.001

port is therefore found for the hypothesis that the responses described as Projection and Rationalization represent different types of perceptual experience.

It is worthy of note that Part I of the experiment involved 10 exposures, whereas Part II involved 34, and that a completely different scene was presented in each slide. The factors of (a) greater unfamiliarity with the situation in Part I than in Part II, (b) changed procedure and changed stimulus material, and (c) different lengths of administration, would all tend to decrease the expected correlations between Projection in Part I and Projection in Part II and between Rationalization in Part I and Rationalization in Part II. Thus a fairly rigorous test of the hypothesis was made, and the values of the coefficients obtained are therefore impressive.

Introspective evidence, as indicated by answers to the questionnaire, gives further support to the main hypothesis. The items and the percentages of answers in the YES, NO, or CANNOT SAY categories are shown in Table 2. All four questions were answered by a majority of the subjects in the direction favoring the hypothesis.

TABLE 2. QUESTIONNAIRE RESULTS

Item	Percent cases answering YES	NO	CANNOT SAY
1. In answering Section III, did you have difficulty in classifying your responses as Type A or Type B?	28	72	—
2. Would you have preferred to rate your responses on a scale, say from 1-5, indicating the *degree* to which your interpretations were placed between the extremes of (A) being made quite readily and (B) being made with considerable effort or delay?	40	50	10
3. Did the nature of your experiences as recorded in Section III seem to fall quite naturally into classes A or B?	66	29	5
4. The record blank is designed so that we may isolate, classify and quantify the functions operating in perception under these conditions. Do you consider that a satisfactory classification is achieved by the record blank?	68	32	—

The majority of those who answered "NO" to Item 4 are the people who would have preferred a rating scale for location of the A and B type experiences, rather than a dichotomy. Absence of difficulty in classifying an experience as A or B (Question 1) does not necessarily imply that the experience "falls quite naturally into classes A or B" (Question 3), since an experience that does not seem naturally to fall into one or other of the categories may logically be classified by the subject in using the formula prescribed by the instructions. It would, however, be consistent with the hypothesis under examination if there was agreement between answers to these two questions. The breakdown is shown in Table 3, which yields a Chi square value of 55.7 significant at $P < .001$.

TABLE 3. AGREEMENT BETWEEN ANSWERS TO QUESTIONS 1 AND 3

		Item 3	
	N = 85	Falls naturally into classes A and B	
		NO	YES
Item 1	Difficulty in classifying as A or B — NO	5	57
	Difficulty in classifying as A or B — YES	21	2

Discussion

It may be concluded from the results of the present experiment that there is good objective evidence to support the view that in perception under less than optimal conditions, two kinds of non-veridical responses occur. The subjective evidence, obtained from introspection by the subjects themselves and reported in answers to the questionnaire, tends to support this view and to strengthen the hypothesis that these different kinds of responses reflect qualitative differences in the nature of the perceptual experiences themselves, rather than differences in the manner of reporting them. This distinction may be important for projective test theory and practice. The types of response discussed recall Rorschach's distinction (5, p. 17) between interpretations in which the *effort* of integration is so great that it is consciously realized as an effort, and interpretations in which subjects are *not* aware of the assimilative effort. A systematic attempt to discriminate Rorschach responses as Type A or Type B as they are defined in the present discussion may lead to improved diagnostic efficiency.

The present results also recall Bruner's affirmation that we know little about "why the Rorschach works" and that a better understanding of this waits upon the day when students of the Rorschach "add to their research something more than 10 ink blots." (1) The present results illustrate his statement that ink blots "are no separate species of perceptual stimulus", and suggest that progress towards a general theory of perception might be fostered by using experimental situations which are, for the subject, a compromise between the extreme freedom of the Rorschach Test and the extreme constraint of the psycho-physical experiment.

Summary

An experiment was conducted in which ambiguous picture material was shown to 89 subjects at gradually increasing levels of illumination.

In Part I, an unconstrained procedure was used, and the protocols were scored by the experimenter in such a way as to discriminate non-veridical responses into two categories, positive "Projection" responses and equivocal "Rationalization" responses. In Part II, the subjects themselves distinguished their Type A (Projection) and Type B (Rationalization) responses. Significant correlations were found between Part I and Part II on the corresponding variables but not between non-corresponding variables. The experimental results, together with introspective data secured from questionnaire answers lend strong support to the view that two types of perceptual experience occur under the conditions used and that these differences cannot be explained away as artifacts of reporting. The data are discussed in relation to projective testing and perceptual theory.

References

1. Bruner, J. S. Perceptual theory and the Rorschach Test. *J. Pers.*, 1948, *17*, 157-168.
2. Jenkin, N. S. An experimental study of visual perception in relation to personality. Unpublished master's thesis. Canterbury University College, New Zealand. 1952.
3. Jenkin, N. S. Some relationships between projective test behavior and perception. *J. clin. Psychol.*, 1955, *11*, 278-281.
4. Postman, L., Bruner, J. S., and McGinnies, E. Personal values as selective factors in perception. *J. abnorm. soc. Psychol.*, 1948, *43*, 142-154.
5. Rorschach, H. *Psychodiagnostics.* New York: Grune and Stratton, 1942.
6. Rubinstein, E. A. Personality integration and the perceptual process. *J. clin. Psychol.*, 1954, *10*, 23-29.

A CONVERSION TABLE FOR RELATING THE MMPI GROUP AND INDIVIDUAL FORM ITEMS*

CHARLES O. GASTON
University of Houston

MAURICE R. SEAQUIST**
Captain, USAF

AND

EARL S. TAULBEE
VA Mental Hygeine Clinic, Omaha, and University of Nebraska College of Medicine

SAUL B. SELLS
USAF School of Aviation Medicine Randolph Field, Texas

This paper presents a conversion table to assist the research worker and clinician in analyzing the item content of the group and individual forms of the Minnesota Multiphasic Personality Inventory[1]. The information presented is not readily available in the publisher's manual. It was prepared to meet needs experienced in our laboratory and is published in response to requests for copies. The chart is useful in relation to the following specific problems:

(1) *To identify the scales for which individual items are scored.* In analyzing item validities with reference to specific criteria, the investigator may be interested in item content in relation to the clinical scales. This information is conveniently presented in the chart, which is keyed both for the Individual and Group Forms.

(2) *To identify corresponding items on the Individual and Group Forms.* Such comparisons are frequently made when comparing clinical with research data, or when pilot studies are made with the Individual Form, and subsequent testing with larger groups utilizes the Group Form.

(3) *To analyze the clinical profiles of patients in terms of specific responses scored on two or more scales.* An inspection of the chart will show that 4 items contribute to 5 scales; 16, to 4 scales; 45, to 3 scales; and 109, to 2 scales. From item 367 to item 566, only 7 items are scored and these, only on the K scale.

(4) *To make a shortened version of the Individual Form.* Instead of scanning the template for each scale to identify the item cards which may be desired; or to discard those which are not scored, the chart may be used to locate the items to be selected from the card deck.

The first column of table 1 is arranged in the item sequence of the Group Form. Column 2 presents the corresponding item code number on the Individual Form. Column 3 indicates the side on which the individual item card is cut, corresponding to "true" or "false" in terms of the Group Form. Column 4 gives the scoring according to the published keys. The last column shows the scales to which the item contributes. Items which are cut on the left side (L) are scored (X) if they are answered "false," and those cut on the right side (R) are scored (X) if answered "true." Certain items, e.g., Group Form item number 6, are scored (0) on certain scales when an (X) is not placed on the Individual Answer Sheet (i.e., left cut answered "true," right cut answered "false.") A few items are scored both ways (XO), in which case, the scales which apply to the (O) scores are printed in bold type. If an investigator should want to obtain the corresponding Group Form items from the Individual Form, he needs only to refer to the chart and fill them in on the Individual Answer Sheet.

REFERENCES

1. HATHAWAY, S. R., and MCKINLEY, J. C. Minnesota Multiphasic Personality Inventory: Manual New York: Psychological Corporation, 1951.

*Editor's Note: This paper is published jointly by two sets of authors as an editorial compromise made necessary by the rare occurrence of almost identical manuscripts being submitted from different sources. Messrs. Gaston and Taulbee had chronological priority by virtue of earlier submission date as of 7 June 1955. The paper by Messrs. Seaquist and Sells, submitted on 2 August 1955, was more complete through inclusion of columns 3, 4 and 5 in table 1. Both sets of authors have agreed to the compromise.

**Now stationed at the University of Texas, Austin, Texas.

TABLE 1. MMPI ITEM CONTENT CHART

Group No.	Card No.	Cut	Score	Scale	Group No.	Card No.	Cut	Score	Scale	Group No.	Card No.	Cut	Score	Scale
(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
1	J-5	L	X	Mf	63	B-20	L	X	Hs	125	B-15	R	X	Hs
2	B-12	L	X	Hs, Hy, D	64	G-37	R	XO	Ma, **D**	126	J-24	R	X	Mf
3	B-27	L	X	Pt, Hy, Hs	65	C-15	L	X	F, Sc	127	H-26	R	X	Pa, Pd, Ma
4	J-34	R	X	Mf	66	H-24	R	X	F	128	H-54	L	X	Hy
5	B-31	R	X	D	67	F-36	R	X	Pt, Pd, D	129	F-33	R	O	K, Hy
6	C-42	R	O	Hy	68	A-50	L	X	Hs	130	B-8	L	XO	Hs, **D**
7	A-55	L	X	Hy, Hs	69	C-50	R	X	Mf	131	H-43	L	X	D
8	F-39	L	X	Pt, Pd, D, Hy, Sc	70	I-45	R	X	Mf	132	J-12	R	X	Mf
9	C-17	L	X	Hy, D, Hs	71	D-48	L	X	K, Hy	133	D-5	L	X	Mf
10	B-11	R	X	Hy, Pt	72	B-17	R	X	Hs	134	G-30	L	XO	K, **Mf, Ma,** Pd
11	B-37	R	X	Ma	73	H-19	R	X	Ma	135	J-45	L	X	L
12	C-43	L	X	Hy	74	J-40	R	X	Mf	136	H-10	R	O	Hy
13	I-34	R	X	Ma, D	75	J-43	L	X	F, L	137	B-48	L	X	Pd, Hy
14	B-19	R	X	F	76	F-49	R	X	Pt, Hy, Sc	138	F-34	R	XO	**K,** D
15	J-41	L	XO	**Pa, Pt, Sc,** L	77	C-45	L	O	Mf	139	G-48	R	X	F
16	G-55	R	X	Pa, Pd, Sc	78	J-18	R	X	Mf	140	J-11	R	X	Mf
17	C-16	L	X	F, Sc	79	F-52	L	X	Mf	141	D-38	L	X	Pd, Hy
18	B-18	L	X	Hs, D	80	I-4	R	O	D, Mf	142	I-37	R	XO	**K,** Pt, D
19	E-16	R	O	Mf	81	I-49	R	O	Mf	143	D-34	R	X	Ma
20	C-47	L	X	Pd, F, Sc	82	I-26	R	O	Pd	144	J-10	R	O	Mf
21	C-6	R	X	Ma, Pd, Sc	83	I-9	L	X	F	145	G-32	R	O	D
22	A-22	R	X	Ma, Pt, Pa, Sc.	84	I-12	R	X	Pd	146	G-33	R	X	F
23	A-15	R	X	F, Hs, Hy, D	85	I-7	R	X	F	147	C-25	R	O	Hy
24	E-37	R	X	Pa, Pd, Sc	86	I-39	R	X	Pt, D	148	C-27	R	O	K, Ma
25	J-32	R	X	Mf	87	I-53	R	X	Mf	149	J-25	R	X	Mf
26	E-3	L	X	Mf, Hy	88	F-38	L	X	D	150	J-55	L	X	L
27	H-25	R	X	Pa, F	89	C-33	R	O	K, Mf, D, Hy	151	H-14	R	X	F, Pa
28	E-9	R	O	Mf	90	J-53	L	X	L	152	B-30	L	X	Pt, D
29	B-16	R	X	Hs	91	E-46	L	X	Pd	153	A-1	L	X	Hs, Hy, D
30	J-51	L	X	K, D, Hy, L	92	J-1	R	X	Mf	154	A-18	L	X	D
31	B-35	R	X	F	93	D-52	R	O	Pa, Hy	155	A-4	L	X	Hs, D, Pd
32	I-27	R	X	Pt, Pd, D, Sc, Hy	94	E-50	R	X	Pt, Pd	156	A-19	R	X	F, Ma, Sc
33	H-12	R	X	Pd, Sc	95	D-10	L	X	D	157	D-33	R	X	Pa, Ma, Sc
34	B-7	R	X	F	96	B-55	L	XO	**K,** Pd	158	F-51	R	X	Pa, D
35	G-53	R	X	Pa, Pd, F, Sc	97	G-47	R	X	Ma, Sc	159	B-2	R	X	D, Pt, Sc
36	A-6	L	X	D, Pt	98	D-18	L	X	D	160	A-3	R	O	Hy, D, K
37	E-17	L	X	Pd, Sc	99	E-34	L	X	Mf	161	A-14	R	X	Hs
38	D-32	R	X	Pd, Sc	100	G-19	R	X	Ma	162	F-25	R	O	Hy
39	G-31	R	O	K, D	101	D-7	L	X	Ma	163	A-5	L	X	Hs, Hy
40	F-1	R	X	F, Sc	102	I-13	R	X	Pd, Pt	164	C-41	L	X	Pt, F
41	F-44	R	X	Pt, D, Sc	103	A-44	L	X	Hy, Hs, Sc	165	J-46	L	X	L
42	B-47	R	X	Pd, F	104	G-7	R	X	D, Sc	166	H-47	R	O	Ma
43	B-28	R	X	D, Hy, Hs	105	J-44	L	X	Ma, L	167	E-15	R	X	Ma
44	A-11	R	X	Hy	106	F-50	R	X	Pt, Pd	168	A-25	R	X	F, Sc
45	J-42	L	X	L	107	G-12	L	X	Hy, D, Pd, Pa	169	H-33	L	X	F
46	A-24	L	X	D	108	A-13	R	X	Hs	170	F-7	L	X	Hy, Pd, K
47	B-1	R	X	Hy, Sc	109	G-50	R	XO	**Pa, Ma, Hy**	171	F-8	R	O	Ma, Pd, K
48	H-22	R	X	F	110	G-54	R	X	Pa, Pd	172	F-9	R	O	Hy
49	D-43	R	X	F	111	D-35	L	X	Pa, Ma	173	C-35	L	X	Pd
50	D-23	R	X	F	112	D-39	L	X	Mf, F	174	A-17	L	X	Hy
51	A-2	L	X	Hs, Hy, D	113	D-42	L	X	F	175	A-16	L	X	Hs, Hy
52	E-24	R	X	D, Sc	114	A-12	R	X	Hs, Hy	176	H-38	L	X	Mf
53	D-21	R	X	F	115	D-17	L	X	F, Mf	177	C-14	L	X	F, Sc
54	E-39	L	X	F	116	D-36	R	O	Mf	178	A-27	L	X	Pt, D, Sc
55	B-9	L	X	Hs, Hy	117	D-50	R	O	Pa, Mf	179	C-51	R	X	Mf, Hy, Sc
56	D-29	R	X	F	118	D-30	R	X	Pd	180	E-44	R	O	Pd, Ma, Hy, K
57	E-36	L	X	D	119	A-37	L	X	Ma, Sc	181	E-31	R	X	Ma
58	D-22	R	O	D	120	J-54	L	X	Mf, Ma, L	182	H-55	R	X	Pt, D, Sc
59	C-24	R	X	Ma	121	H-11	R	X	F, Pa, Sc	183	D-51	R	O	K, Pd
60	J-52	L	X	L	122	A-23	L	X	D, Pt	184	H-21	R	X	F
61	G-4	R	X	Pd	123	H-4	R	X	F, Pa	185	A-35	L	X	F
62	A-47	R	X	Hs	124	D-54	R	O	Pa, K, Hy	186	A-45	R	X	Hy

TABLE 1. MMPI ITEM CONTENT CHART (continued)

Group No.	Card No.	Cut	Score	Scale	Group No.	Card No.	Cut	Score	Scale	Group No.	Card No.	Cut	Score	Scale
(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
187	A-43	L	XO	**Mf**, Sc	250	E-7	R	X	**Ma**	313	D-55	R	O	Pa
188	A-32	L	X	Hy, Hs	251	A-21	R	X	Ma, Sc	314	J-41	L	XO	**Pa, Pt, Sc,** L
189	A-40	R	X	Pt, D, Hs, Hy	252	I-8	R	X	F	315	G-55	R	X	Pa, Pd, Sc
190	A-10	L	X	Hs, Hy	253	E-54	L	O	Hy	316	D-53	R	O	Pa, K
191	B-4	R	O	D	254	E-33	L	X	Mf	317	F-10	R	X	Pt, Pa
192	A-42	L	X	Hy, Hs, Sc	255	J-49	L	X	L	318	F-39	L	X	Pt, Pd, D, Hy, Sc
193	B-6	L	O	D	256	C-46	R	X	F	319	D-46	R	O	Pa
194	A-20	R	X	Ma, Sc	257	C-29	L	X	F	320	C-52	R	X	Sc
195	J-47	L	X	L	258	D-15	L	X	F	321	F-4	R	X	Pt
196	H-27	L	X	Sc, F	259	F-42	R	X	D, Sc	322	C-18	R	O	K, Sc
197	H-15	R	X	F	260	C-36	R	O	Mf	323	H-12	R	X	Pd, Sc
198	F-15	L	X	Mf	261	I-44	R	X	Mf	324	C-11	R	X	Sc
199	D-6	L	X	F	262	E-49	L	X	Mf	325	C-5	R	X	Sc
200	H-17	R	X	F	263	B-3	R	XO	**D, Ma**	326	A-22	R	X	Ma, Pt, Pa, Sc
201	F-5	R	O	Hy, Pd	264	I-40	L	X	Mf	327	B-54	L	X	Pa
202	G-9	R	X	F, Pa, Sc	265	E-53	R	O	Hy	328	I-27	R	X	Pt, Pd, D, Sc, Hy
203	J-19	R	X	Mf	266	G-20	R	X	Pt, Sc, Ma	329	B-33	R	O	Pt
204	J-23	R	X	Mf	267	E-43	R	O	K, Pd, Ma, Hy	330	A-41	L	X	Sc
205	G-44	R	X	F	268	G-16	L	XO	Pa, **Ma**	331	G-53	R	X	Pa, Pd, F, Sc
206	D-8	R	X	F	269	I-1	R	X	F	332	A-38	R	X	Sc
207	E-41	L	X	D	270	G-38	L	X	D	333	E-37	R	X	Pa, Pd, Sc
208	C-48	R	O	D	271	D-49	R	XO	Ma, **D**	334	A-30	R	X	Sc
209	G-3	R	X	F	272	G-23	L	X	F, K, D	335	A-26	R	X	Sc
210	G-11	R	X	F, Sc	273	A-48	R	X	Hs, Sc	336	F-31	R	X	Pt
211	B-32	R	X	F	274	A-31	L	X	Hy, Hs	337	G-1	R	X	Pt
212	B-50	R	X	Ma, Sc	275	H-7	R	X	F, Pa	338	G-6	R	X	Pa
213	G-40	R	O	Mf, Hy	276	E-20	L	X	F, Sc	339	G-10	R	X	Sc
214	B-25	L	X	Mf	277	E-11	R	X	Ma	340	G-22	R	X	Pt
215	B-42	R	X	Pd, F	278	H-2	R	X	Mf	341	G-34	R	X	Pa, Sc
216	B-49	R	X	Pd	279	B-40	R	XO	Ma, **Hy**	342	G-35	R	X	Pt
217	F-46	R	XO	**K,** Pt, Mf	280	D-47	R	O	Mf	343	G-36	R	X	Pt
218	I-6	R	X	F	281	A-36	L	X	Pa, Hs, Sc	344	G-45	R	X	Pt
219	J-7	R	O	Mf	282	C-10	R	X	Mf, Sc	345	G-51	R	X	Sc
220	C-13	L	X	F, Sc	283	J-29	R	O	Mf	346	G-41	R	X	Pt
221	I-50	L	X	Mf	284	H-3	R	X	Pd, Pa	347	G-52	L	X	Pa
222	F-14	R	X	Ma	285	J-50	L	X	D, L	348	H-9	L	X	Pa
223	I-51	R	O	Mf	286	E-21	R	X	F	349	H-13	R	X	Pt, Sc
224	B-52	R	X	Pd	287	H-31	L	X	Pd	350	H 20	R	X	Sc
225	J-48	L	X	L	288	A-15	R	X	F, Hs, Hy, D	351	H-28	R	X	Pt
226	C-3	R	X	Ma, Mf	289	D-44	L	X	Ma, Hy, Pd	352	H-32	R	X	Pt, Sc
227	B-38	R	X	F	290	I-34	R	X	Ma, D	353	H-52	L	X	Pt
228	G-26	L	O	Ma	291	H-8	R	X	F, Pa, Sc	354	H-53	R	X	Sc
229	E-35	L	X	Mf	292	E-23	R	O	Hy	355	I-2	R	X	Sc
230	B-10	L	X	Hs, Hy	293	H-6	R	X	F, Pa	356	I-23	R	X	Pt, Sc
231	D-2	R	XO	Mf, **Pd**	294	E-12	L	X	Pa, Pd	357	I-25	R	X	Pt
232	D-26	R	X	Ma	295	J-26	R	X	Mf	358	G-43	R	X	Pt
233	C-31	R	XO	Ma, **D**	296	G-18	L	X	D, Pd, K	359	G-42	R	X	Pt
234	G-29	R	O	K, Hy	297	C-53	R	X	Mf, Sc	360	H-30	R	X	Pt, Sc
235	B-51	L	X	Pd	298	E-1	R	X	Ma	361	I-32	R	X	Pt
236	F-45	R	X	D	299	A-9	R	X	Mf, Pa	362	F-10	R	X	Pt, Pa
237	C-7	L	X	Pd	300	J-31	L	X	Mf	363	I-3	R	X	Sc
238	G-21	R	X	Pt, Sc, Hy, Ma	301	I-35	R	X	Pt, Sc	364	H-16	R	X	Pa, Sc
239	C-12	R	X	Mf, Pd	302	E-17	L	X	Pd, Sc	365	H-49	R	X	Pa
240	E-48	R	X	Ma	303	F-18	R	X	Sc	366	F-35	R	X	Pa, Pt, Sc
241	B-36	R	XO	**D,** Sc	304	F-3	R	X	Pt	367	H-36	L		
242	A-8	L	X	D	305	F-35	R	X	Pa, Pt, Sc	368	F-13	R		
243	A-46	L	X	Hy, Hs	306	E-40	L	X	Sc	369	D-9	L		
244	I-30	R	X	Pd	307	E-42	R	X	Sc	370	C-21	L		
245	B-53	R	X	Pd, F	308	C-6	R	X	Ma, Pd, Sc	371	F-6	L		
246	A-54	R	X	F	309	E-38	L	X	Sc	372	G-14	L		
247	C-9	R	X	F	310	C-47	L	X	Pd, F, Sc	373	D-13	R		
248	G-25	R	O	Pd, D	311	D-32	R	X	Pd, Sc	374	F-43	R	O	K
249	D-16	L	X	Mf	312	E-22	R	X	Sc	375	C-34	R		

(Table continued on next page.)

TABLE 1. MMPI ITEM CONTENT CHART (continued)

Group No.	Card No.	Cut	Score	Scale	Group No.	Card No.	Cut	Score	Scale	Group No.	Card No.	Cut	Score	Scale
(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
376	E-13	L			440	E-47	L			504	F-21	R		
377	E-27	R			441	J-22	R			505	G-24	R		
378	E-51	L			442	B-29	L			506	I-33	R		
379	F-48	L			443	C-19	R			507	I-16	R		
380	G-49	L			444	F-29	L			508	A-29	L		
381	G-27	R			445	D-31	L			509	F-12	R		
382	I-17	R			446	D-37	R			510	H-42	R		
383	E-52	R	O	K	447	D-40	R			511	F-16	R		
384	F-17	R			448	H-1	R			512	J-9	R		
385	H-34	R			449	E-29	L			513	J-17	L		
386	G-17	R			450	E-30	L			514	J-38	R		
387	D-20	L			451	E-32	L			515	B-46	L		
388	H-46	R			452	E-45	R			516	C-2	L		
389	I-15	R			453	E-19	R			517	F-40	R		
390	F-53	R			454	E-25	R			518	G-5	R		
391	E-28	L			455	E-26	R			519	B-26	R		
392	H-35	R			456	E-8	R			520	F-11	L		
393	I-5	L			457	B-43	R			521	F-2	L		
394	I-14	L			458	J-39	R			522	H-40	L		
395	I-10	R			459	B-44	R			523	A-52	L		
396	F-47	R			460	B-41	L			524	H-44	L		
397	I-31	R	O	K	461	C-28	L	X	K	525	H-37	R		
398	I-38	R	O	K	462	B-22	L			526	F-55	R		
399	G-28	L			463	J-37	R			527	C-1	L		
400	H-18	R			464	H-23	L			528	A 51	L		
401	H-41	L			465	C-22	R			529	J-35	L		
402	I-18	R			466	B-39	L			530	A-53	R		
403	I-11	L			467	G-39	R			531	I-20	R		
404	F-28	R			468	F-32	L			532	J-16	L		
405	A-39	L			469	F-30	R			533	B-14	L		
406	I-22	L	X	K	470	C-54	R			534	I-24	L		
407	F-37	L			471	D-28	R			535	B-5	R		
408	F-27	L			472	G-46	R			536	I-48	R		
409	I-19	L			473	H-51	R			537	I-52	R		
410	E-10	R			474	B-23	L			538	I-42	R		
411	I-29	R			475	E-4	R			539	H-39	L		
412	A-7	L			476	D-24	R			540	A-34	L		
413	G-8	R			477	E-5	R			541	A-49	R		
414	F-54	R			478	E-14	L			542	B-21	L		
415	I-28	L			479	E-18	L			543	G 2	R		
416	C-23	R			480	J-36	R			544	J-13	R		
417	F-24	R			481	I-21	R			545	J-15	R		
418	I-36	R			482	E-55	L			546	C-39	L		
419	D-27	R			483	D-19	L			547	J-33	L		
420	D-25	R			484	B-45	R			548	D-1	L		
421	C-8	R			485	D-4	R			549	I-41	R		
422	C-4	R			486	B-24	L			550	I-46	L		
423	I-47	L			487	F-41	R			551	H-5	R		
424	B-13	R			488	D-11	L			552	C-37	L		
425	B-34	L			489	F-22	R			553	H 48	R		
426	F-26	R			490	D-12	R			554	J-20	R		
427	C-55	R			491	D-14	L			555	J-4	R		
428	C-38	L			492	H-45	R			556	J-14	R		
429	C-40	L			493	C-30	L			557	J-8	R		
430	C-49	L			494	H-50	R			558	D-3	L		
431	I-54	R			495	C-32	L			559	J-6	R		
432	D-41	R			496	A-33	L			560	A-28	R		
433	J-21	R			497	C-44	L			561	J-28	L		
434	J-2	R			498	F-19	L			562	J-30	L		
435	J-3	R			499	H-29	R			563	I-43	L		
436	D-45	L			500	G-15	R			564	C-20	R		
437	E-6	R			501	C-26	L			565	I-55	R		
438	E-2	R			502	F-20	L	X	K	566	J-27	L		
439	G-13	R			503	F-23	R							

GROUP PSYCHOTHERAPY IN CONJUNCTION WITH A PHYSICAL TREATMENT

RICHARD C. COWDEN, MELVIN ZAX AND J. A. SPROLES[1]

VA Hospital, Gulfport, Mississippi

PROBLEM

In most of the appraisals of the effectiveness of psychotherapy there lies the implication that it is extremely difficult to evaluate[4, 5, 7, 9]. Most evaluations are made in terms of the therapist rating the patient as improved, cured, or no different. Two factors enter to negate many ratings of this type. One is in the definition of these categories. The other is a personal bias on the part of the rater. That there may be some personal subjective involvement on the part of the therapist in making such a rating is accepted almost without question and has recently been experimentally verified[8].

With chronic schizophrenic patients this evaluation becomes doubly difficult as does the therapy itself. These patients are frequently highly inaccessible to psychotherapy and require a long period of time to establish any type of rapport with them as was shown by the Powdermaker and Frank study of group psychotherapy[6]. There has long been a need for some technique to make these patients more accessible. With the advent of the new alleviating drugs, chlorpromazine and reserpine, there appeared a means to do this. These drugs promise the benefits of present pharmaceutical treatments without many of the liabilities. Neither drug requires constant supervision or leaves the patient in a comatose state for a long period of time and as a result less accessible than before administration of the drug. Other physical treatments quiet the patient temporarily but prevent any immediate further treatment in the form of psychotherapy. Cowden et al[2, 3] have shown the efficacy of chlorpromazine when used alone and in conjunction with group psychotherapy. This experiment uses the Rauwolfia Serpentina extract, reserpine. Reserpine produces a tranquilizing and sedative effect, a gradual sustained lowering of blood pressure, and a slowing of the heart rate. It acts chiefly in the region of the hypothalamus and certain medullary centers. It does not in any way alter the schizophrenic process itself.

This study sought to determine the effects of group psychotherapy, alone and in conjunction with reserpine, on chronic severely disturbed and agitated schizophrenic patients. At the same time it measured the effectiveness of the drug, reserpine, alone. It also attempted to define objective means of measuring the changes wrought by these treatments.

METHOD

Thirty-two chronic, severely disturbed and agitated schizophrenic patients residing on the maximum security ward were selected for the experiment. They were assigned to one of four groups of eight patients by the ward psychiatrist. A hospital employee, unfamiliar with the patients, then decided which group would get which type of treatment. Group A received group psychotherapy and reserpine; Group B received group psychotherapy and a placebo (sodium bicarbonate); Group C received only reserpine; and Group D remained as usual, serving as the control group. In the course of the experiment one patient was lost from Group C for refusing to take the drug. After being placed in one of the four groups, each patient was then rated by two psychiatrists on the Lorr Multidimensional Scale for Rating Psychiatric Patients. Each patient was also rated by two ward attendants on a seven-point behavioral rating scale devised for this purpose and independently checked for reliability[3]. All patients were given the following psychological tests: Bender-Gestalt, H-T-P, H-T-P chromatic, self-concept drawing, a sentence completion test,

[1]The authors wish to thank John Bowland, psychology trainee, University of Florida, for serving as co-therapist.

and a modified form of the TAT. This testing was repeated at the completion of the experiment six months later.

The data referring to age, behavioral rating, and length of hospitalization for each group are summarized in Table 1. All of the patients had previously received

TABLE 1. MEAN AGE, BEHAVIORAL RATING, AND LENGTH OF HOSPITALIZATION OF SUBJECTS.

Group	Age	Behavioral Rating	Hospitalization in months
A	38.5	2.7	54
B	31.9	3.4	50
C	38.1	3.3	60
D	33.9	2.4	71

insulin coma therapy and electroshock therapy with no remarkable improvement. In each group there was one patient who had received a lobotomy. All of these patients were considered violent and dangerous to themselves or others. Medication was started on August 23 and ended on February 25. The patients were started on one mg. per day and dosage was adjusted thereafter according to the behavior of the patient. Maximum dosage was eight mg. per day; average dosage was four mg. per day at the completion of the experiment. There have been no remarkable side effects. All the patients had blood studies with negative results.

The group psychotherapy sessions were held three times per week, one hour per session. There were two therapists, a staff psychologist and a clinical psychology trainee. The therapy was aimed primarily at inducing behavioral changes. It attempted to get the patients to express themselves in a way that was personally satisfying and socially acceptable. The methods used to obtain the changes varied from being very directive and at times punitive to being completely accepting and nondirective. This varied according to the individual patient and the approach depended on determining the individual needs, defenses, and stresses operating on the patient. The most difficult aspect of therapy with this type of patient was in getting them to function as a group; there was a marked tendency for each patient to try to dominate the attention of both therapists. This rivalry was taken advantage of and led to the patients communicating with each other as well as with the therapists. At no time was any assaultive behavior permitted while therapy was in session; this limit was firmly impressed on the patients since such behavior was not uncommon. Interpretations were made very sparingly and then they dealt with immediate behavior and feeling. No attempt was made to interpret in a dynamic or historical sense or to provide insight to the patient.

The measurements of improvement were purely behavioral ratings, only one of which required the subjective clinical judgment of any of the professional staff members. They were as follows:

1. Ratings by two staff psychiatrists on the Lorr Multidimensional Scale.
2. Ratings by two ward attendants on a seven-point behavioral rating scale. These ratings were made one week before medication began and again on the day medication was stopped six months later.
3. Decline in number of neutral wet packs.
4. Decline in number of times placed in room seclusion.
5. Decline in number of fights engaged in.
6. Decline in number of disturbed reports made by the nurses.
7. Transfer to a better ward requiring a higher level of integration.
8. Discharge from the hospital.

The number of packs, room seclusions, fights, and disturbed reports was determined for the six months prior to treatment and compared with the six months during which the patients received treatment. All of the above are objective behavioral measures since no dramatic changes in personality structure were expected to

occur over a period of six months. The original ratings of the patients by the two psychiatrists on the Lorr Multidimensional Scale were so divergent that they were eliminated from further consideration.

Results

The differences between the pre- and post- behavioral scale ratings by the two ward attendants were first considered. The rating used for tabulation was the mean of the ratings by the two attendants. A parametric analysis of variance for the four groups yielded an F value of .49 for 3 and 24 degrees of freedom, which is not significant. The Kruskall-Wallis test, a nonparametric analysis of variance, gave a chi-square of 3.18 and for three degrees of freedom was significant at between the .15 and .10 point. Thus, although the *degree* of change from group to group was not significant, the nonparametric anova strongly suggests that the direction of the change was more consistent in some groups. Of the six parametric *t*-tests between all groups, only one approached significance. This was a *t*-value of 1.66 for the difference between the therapy-plus-drug group and the control group, which for 12 degrees of freedom has a P-value between the .10 and .05 point.

In analyzing the data referring to the number of packs, room seclusions, fights, and disturbed reports, Wilcoxon's nonparametric *t*-test for paired replicates was used along with a chi-square to compare the six months before treatment to the six months following its introduction. These results as presented in Table 2 appear quite spectacular. Group A showed a significant improvement on all of the measures, while the control group (Group D) remained practically the same. The other two groups showed significant changes also but not in all areas. It is to be noted here that the Wilcoxon *t*-test, which is not influenced by extreme cases, does not substantiate chi-square in all measures. In these cases the significant drop is due to dramatic changes in one or two members of a group. This is discussed later in greater detail. That there did occur definite behavioral changes is more than substantiated by Table 2. Fisher's exact method was used to determine the chi-square for test differences between the four groups for each of the seven psychological tests. None of the differences on any of the tests approached significance.

Table 2. Behavioral Changes as a Result of Treatment

Group	Measure	Pre	Post	X^2	Wilcoxon R
(A) Reserpine and Psychotherapy	Packs	53	16	19.8**	3.0*
	Seclusion	80	40	13.2**	——
	Fights	30	14	5.8*	1.5**
	Dist. Reports	45	6	29.8**	1.5**
(B) Psychotherapy and Placebo	Packs	56	15	23.6**	10.5
	Seclusion	1	0	——	——
	Fights	18	7	4.8*	0.5**
	Dist. Reports	18	1	15.2**	0.5**
(C) Reserpine Only	Packs	48	11	23.2**	7.5
	Seclusion	0	0	——	——
	Fights	13	8	1.2	6.0
	Dist. Reports	39	10	17.6**	0.5**
(D) Control	Packs	73	80	——	8.5
	Seclusion	83	84	——	——
	Fights	30	28	——	16.0
	Dist. Reports	26	21	——	17.0

**P-value less than .01
*P-value less than .025

Discussion

Alexander[1, p. 410] states that "psychotherapy may be greatly and at times immeasurably facilitated" by physical treatment. The results of this study tend essen-

tially to confirm this. Also, they are in complete agreement with a similar study which used the drug chlorpromazine[3], the primary difference between the two studies being the type of patient used. Although there appeared to be no dramatic personality change which should have been reflected in the psychological tests and the behavioral rating scale, the fact remains that marked behavioral changes occurred as revealed by the data in Table 2. Although none of these patients improved sufficiently to warrant discharge, some did show enough change to warrant transfer to better wards that required more personal responsibility and integration along with additional privileges. Three men from Group A were transferred and one each from Groups B and C. At the completion of a two-month follow-up period of observation these patients were maintaining their improvement.

In all measures the greatest improvement was made by the group which received reserpine in addition to group psychotherapy. It is very likely that psychotherapy without the drug would have been extremely difficult and more time consuming. It would have required a much longer period of time to achieve the first goal of therapy, that of making the patient comfortable in the group. This was observed by the therapists in that at the conclusion of the experiment the drug-plus-therapy group was functioning more as a unit than the group which did not receive the drug.

When one keeps in mind the fact that these patients had resided on the maximum security ward for a long period of time, as much as eight years, with very little formal treatment of any type, the behavioral change is quite remarkable. Further analysis of Table 2 suggests the source and reason for this change. It is noteworthy that there is a discrepancy between the chi-square and the Wilcoxon R. The chi-square is a measure of the difference in the *actual number* of incidents occurring whether this number is a function of only one patient or eight patients. The Wilcoxon *t*-test on the other hand, is not influenced by the extreme cases but rather is a measure of the consistency of the direction of change within each group. Where the two statistics are at variance the significant chi-square is a function of a marked change in one or two patients who were relatively more agitated and anxious. In all the groups, the greatest change occurred with those patients who showed the highest level of manifest anxiety, agitation, and a tendency to "act out" their impulses.

Summary

Four groups of eight severely disturbed, chronic schizophrenic patients were selected by the ward psychiatrist. One group received group psychotherapy and the new drug reserpine; another received psychotherapy and a placebo; the third group received only the drug reserpine and the fourth acted as a control group. After six months of treatment, behavioral changes were measured. No significant improvement occurred on the psychological tests. Marked improvement occurred with the behavioral measures. The greatest improvement took place when reserpine was used along with group psychotherapy.

References

1. Alexander, L. *Treatment of mental disorder.* Philadelphia: Saunders, 1953.
2. Cowden, R. C., Zax, M. and Finney, R. C. A preliminary note on the use of chlorpromazine with neuropsychiatric disorders. *Arch. Neurol. Psychiat.*, in press.
3. Cowden, R. C., Zax, M., Hague, J. R. and Finney, R. C. Chlorpromazine: Alone and as an adjunct to group psychotherapy in the treatment of psychiatric patients. *Amer. J. Psychiat.*, in press.
4. Eysenck, H. J. The effects of psychotherapy: an evaluation. *J. consult. Psychol.*, 1952, 16, 319-324.
5. Luborsky, L. A note on Eysenck's article "The effects of psychotherapy: an evaluation." *Brit. J. Psychol.*, 1954, 45, 129-131.
6. Powdermaker, F. B. and Frank, J. D. *Group psychotherapy.* Cambridge, Mass.: Harvard University Press, 1953.
7. Rosenzweig, S. A transvaluation of psychotherapy—a reply to Hans Eysenck. *J. abn. soc. Psychol.*, 1954, 49, 298-304.
8. Saslow, G. Psychotherapy. In *Annual Review of Psychology*, 1954, 5, 311-336.
9. Zubin, J. and Windle, C. Psychological prognosis of outcome in the mental disorders. *J. abn. Psychol.*, 1954, 49, 272-282.

SCORE EQUIVALENCE OF THE WECHSLER-BELLEVUE* INTELLIGENCE SCALES, FORMS I AND II

JOHN R. BARRY AND ALBERT L. KUBALA
SAMUEL C. FULKERSON MAURICE R. SEAQUIST

USAF School of Aviation Medicine, Randolph Field, Texas

Introduction

In evaluating the equivalence of Form I and Form II of the Wechsler-Bellevue Adult Intelligence Scale, three somewhat different questions are appropriate: (a) Do subjects within a given sample retain the same rank on both forms, i.e., is there test reliability? (b) How strong is the transfer of training, or practice effect? (c) Practice effects aside, do subjects get equivalent scores on the two forms?

Several investigators have explored the question of reliability, using the two tests as equivalent forms. Cronbach[1] presented some previously unpublished correlations computed by Wechsler between subtests of Forms I and II for 100 patients with IQ's of 40 to 100. The coefficients ranged from .34 to .81, with the Picture Completion and Comprehension subtests showing the lowest reliabilities, and the Digit Symbol and Block Design subtests showing the highest. Gibby[5] administered both forms in a single session to 32 psychoneurotic out-patients, who ranged in age from 20 to 45 and who had a mean IQ of 108 $\pm$ 14. The form administered first was alternated. Gibby found correlations between the subtests of the two forms ranging from .20 to .93, with the Comprehension and Picture Arrangement subtests having the lowest coefficients and the Vocabulary, Picture Completion and Block Design subtests, the highest. The over-all agreement between the coefficients reported by Cronbach and Gibby is not high.

In his manual for Form II, Wechsler[7] stated that because of the close similarity between the two forms, the diagnostic signs for the Form I subtests probably could be applied to Form II subtests without serious error. However, Gibby reported that the test profiles on the two forms for the same individual showed little similarity. Nevertheless, he concluded that the Verbal, Performance, and Full Scale IQ's were correlated highly enough to be used interchangeably in the clinic.

Gerboth[4] has reported an interaction between practice effect and form in a study in which both forms were administered to 100 Washington University students of both sexes, ranging in age from 17 to 24. The mean IQ was 124 $\pm$ 6. Form I was given initially to half the group and Form II, initially to the other half. The interval between testing was approximately four months. When Form II was administered first, the mean Form II IQ was about five points lower than the Form I IQ. No such practice effects were found when Form I was administered first. Correlations between subtest scores also were found to be lower when Form II was given first.

The differences among the findings described above indicate that score equivalence between Form I and Form II has not been unequivocally established. The differences may be due in part to variations in sampling and testing conditions, but these differences are no greater than are generally met in practice. This study represents a further exploration of score equivalence with an attempt to evaluate both the agreement between Form I and Form II scores and the practice effects over a 24-hour period. It is based on a sample which is restricted in IQ range, but which is representative of the cases met in aeromedical practice where officers of above average intellectual endowment are examined.

Method

The present research sample consists of 38 flyers who were referred for consultation concerning psychological factors related to inefficiency in flying duties. They

*Appreciation is expressed to Drs. S. B. Sells and D. K. Trites, who assisted materially in this research.

were referred for evaluation during the first six months of 1954. They ranged in age from 20 to 36, and reported from 12 to 17 years of education. All were tested on successive days by experienced examiners, the form administered first being alternated from patient to patient. These patients are believed to be representative of the intellectually homogeneous group of flying personnel who comprise the principal category of patients referred for consultation in the Air Force for mildly incapacitating nervous complaints.

Nineteen patients received Form I followed by Form II, usually on the next working day. The other 19 patients received the two forms in the opposite order. An analysis of variance was made for the scores on each of the subtests, except Vocabulary, and for the Verbal, Performance, and Full Scale IQ's. Thus, 13 separate form-by-order-by-subject analyses of variance were made.[1] Three questions were formulated: (a) Do the two forms yield significantly different scores? (b) Are there significant practice effects? (c) Is practice effect a function of the form which is presented first?[2] In addition to these analyses, the equivalence of each individual's subtest scores on the two forms was compared by correlation. It was suspected that the person-to-person subtest score equivalence might be low, even when the group means were not significantly different. Finally, in order to investigate this equivalence further, frequency distributions of the differences between Form I and Form II IQ's were made. These distributions allowed an examination of the person-to-person differences between the same scores on different forms.

Results

Practice Effects. The three IQ's and four of the subtests showed significant practice effects. There were no differences in practice effect associated with the order of test administration. Group means for the three IQ's and the four subtests which showed practice effects are presented in Table 1. The F ratios from the analyses of

Table 1. Significant Effects of Practice on Alternate Wechsler-Bellevue Administrations.

Scores	Form I administered first		Form II administered first		Form by order F ratio*
	Form I Mean	Form II Mean	Form II Mean	Form I Mean	
Digit Span	11.32	12.68	10.95	11.79	6.93
Arithmetic	13.11	13.42	12.89	14.84	6.47
Block Design	13.79	14.26	14.16	15.00	6.79
Digit Symbol	11.16	12.68	13.00	13.53	39.55
Verbal IQ	119.53	121.84	118.63	120.53	4.94
Performance IQ	121.37	124.16	122.58	126.84	8.92
Full Scale IQ	122.32	126.11	123.79	125.89	12.16

*The form by order analysis of variance was evaluated using as an error term the subjects' by form variance, df 1 and 36. An F ratio of 4.11 is significant at the 5 percent level, and 7.39 at the 1 percent level in this table.

variance listed in the last column of this table suggest that the mean differences between the forms within each order were greater than what would be expected by chance. The actual amount of gain in test score with practice can be determined by comparing the group means on the first administered forms with those on the second administered forms. For the subtest scores and IQ's listed in Table 1 significant

[1]The design employed for these analyses was very similar to that described by Edwards[2] on page 295.

[2]For example, Gerboth[4] found that Form II, when given first, had a greater facilitating effect on Form I scores than Form I had on Form II scores when the order of administration was reversed.

practice effects were found regardless of which form was taken first.[3] This finding is somewhat different from that of Gerboth, cited above. It should be noted that Bartlett's test for homogeneity of variance revealed significant heterogeneity ($p < .05$) for the Performance IQ's; hence, the interpretation of this F ratio is less certain.

TABLE 2. A COMPARISON OF THE FORMS I AND II SCORES. (N = 38)

Test score	Form I Mean	Form I S.D.	Form II Mean	Form II S.D.	Form I - Form II Mean Difference	Form I - Form II Pearson r
Information	12.76	1.17	12.74	1.25+	.02	.48*
Comprehension	13.26	2.28	13.03	2.39	.23	.42*
Digit Span	11.55	3.06	11.82	3.07	−.27	.60*
Arithmetic	13.97	3.26	13.16	1.86	.81	.47*
Similarities	12.84	1.94	11.95	1.88	.89**	.42*
Picture Arrangement	12.95	2.95	12.87	2.36	.08	.16
Picture Completion	13.89	1.12	11.97	1.46	1.92*	.10
Block Design	14.39	1.75	14.21	1.98	.18	.63*
Object Assembly	12.79	1.71	13.95	1.69	−1.14*	.38**
Digit Symbol	12.34	2.40	12.84	2.17	−.50**	.70*
Verbal IQ	120.03	7.49	120.24	6.67	−.21	.66*
Performance IQ	124.11	9.45	123.37	8.91	.74	.62*
Full Scale IQ	124.11	7.44	124.95	6.64	−.84	.66*

* = $p < .01$
** = $p < .05$

Score Equivalence. The Form I and Form II mean scores are listed in Table 2. The analysis of variance indicates that differences between the scores on four of the subtests were significant beyond the five per cent level of confidence. Of these, the Similarities and Picture Completion subtests gave higher scores on Form I and the Object Assembly and Digit Symbol subtest scores were higher on Form II. These subtests clearly do not yield equivalent scores. From the last column of Table 2 it can be seen that three of these four subtests have statistically significant correlations when the two forms are compared. Thus, the patients tended to retain the same rank on both forms, although the forms did not appear equivalent. However, the Picture Completion subtest does not have a significant form-to-form correlation. Differences in scores on this subtest apparently involve differences in the composition of the items.

Intercorrelations between Forms I and II for the present study are listed in the last column of Table 2. It should be noted that the person-to-person correspondence for some of the subtest scores is very low, although the group means for these scores are quite comparable. For example, on the Picture Arrangement subtest the Form I and II means are 12.95 ± 2.95 and 12.87 ± 2.36, but the correlation is low ($r = .16$). Thus, one could not predict accurately one subtest score from the other score in this restricted and homogeneous sample.

It has been suggested from other studies[1, 3, 4, 7] that the Form I and II IQ's were correlated to such an extent that they could be used interchangeably in the clinic. However, the present study suggests that a large amount of the variance associated with scores from one form can *not* be predicted from a knowledge of the scores of the other form. In order to investigate this further, an analysis was made of the differences in IQ's from frequency distributions of these differences. These distributions were made separately for the two groups (those who received Form I first and those who received Form II first).

[3]Since each of the subtests showing "practice effects" is a speeded test, it is possible that a part of the observed increase in speed of performance was due to a reduction in task-induced anxiety on the second test.

For all three intercorrelations among the IQ's, the standard error of estimate was approximately five IQ points. For this reason and in accordance with the clinical judgment of the authors, a practically meaningful difference between the Form I and Form II IQ's was defined as six or more IQ points (i.e., greater than five). In addition, a correction for the practice effects shown in Table 1 was made. Specifically for the Full Scale IQ differences, a correction of four IQ points was applied in the group which received Form I first and of two points in the other group. This correction was intended to compensate for the practice effects due to taking the tests on successive days. Of the 19 patients who received Form I first, four had differences of six or greater between the Form I and Form II IQ's (after the correction was made). Of the 19 who received Form II first, five had this large a difference between their total IQ's on the two forms. Considering both groups together, 10 and 16 patients had differences this large between their Verbal and Performance IQ's, respectively. Although the Verbal and Full Scale IQ's for Forms I and II were more equivalent than the Performance IQ's, the implication of this analysis is that the forms should be used interchangeably only with great caution.

Conclusions

1. Practice effects, regardless of which form was administered first, were found for four subtests and for the three IQ scales on the Wechsler-Bellevue Intelligence Scales, Forms I and II.

2. Although the Form I and Form II Verbal, Performance and Full Scale IQ's were significantly correlated, a further study of the data suggests that the scores should be used interchangeably only with great caution, and after corrections for practice effects are made. It must be remembered that such corrections probably will vary with the size of the interval between tests.

3. The individual subtest scores were not equivalent from form to form. The Form II Object Assembly and Digit Symbol subtests yielded higher scores (possibly were easier) than the corresponding Form I subtests; the Form II Similarities subtest, lower scores. The person-to-person score equivalence on the Picture Completion and Picture Arrangement subtests was negligible (correlations of .10 and .16). Thus, the application of Form I subtest patterns to Form II subtests should be questioned.

References

1. Cronbach, L. J. *Essentials of Psychological Testing.* New York: Harper and Brothers, 1949.
2. Edwards, A. L. *Experimental Design in Psychological Research.* New York: Rinehart, 1950.
3. Fisher, R. A. *Statistical Methods for Research Workers.* Edinburgh: Oliver & Boyd, 1936.
4. Gerboth, R. A study of two forms of the Wechsler-Bellevue Intelligence Scale. *J. Consult. Psychol.*, 1950, *14*, 365-370.
5. Gibby, R. G. A preliminary survey of certain aspects of Form II of the Wechsler-Bellevue Scale as compared to Form I. *J. Clin. Psychol.*, 1949, *5*, 165-169.
6. Patterson, C. H. *The Wechsler-Bellevue Scales. A Guide for Counselors.* Charles C. Thomas, Springfield, Illinois, 1953, pp. 14-20.
7. Wechsler, D. *The Wechsler-Bellevue Intelligence Scale, Form II. Manual.* New York: Psychological Corporation, 1946.

CALL FOR PAPERS

This Journal is now able to offer immediate publication for high priority research reports. Increased efficiency of publication processes now permits publication of acceptable papers in three months from date of receipt. Submit your studies now.

UNREALISTIC REPORTING AS A FUNCTION OF EXTRAVERTED NEUROSIS[1]

J. D. KEEHN

American University of Beirut
Beirut, Lebanon

PROBLEM

One of the factors that Mowrer particularly stresses in the development of neurosis is that of cheating or the violation of the 'canons of realistic reporting' in order to avoid punishment[9]. Attention to this point arises from his view of the neurotic as super-ego deficient in contradiction to the orthodox Freudian theory of neurosis as a function of over-severity of the super-ego.

Eysenck[5], following his postulation of the independent personality factors of neuroticism and introversion-extraversion[3], has recently attempted to reconcile the views of Freud and Mowrer in terms of Hull's[8] postulate of reactive inhibition, arguing that while Freud's point of view is compatible with the anxiety syndrome, Mowrer's is more pertinent to the hysterical disorders. Thus Eysenck has presented evidence that these opposing theories of *neurosis* are concerned with *introversion* and *extraversion* respectively. By so equating Mowrer's analysis with hysteria, or extraverted neurosis, it would be expected that Eysenck would predict that not all neurotics but only hysterics would exhibit 'unrealistic reporting'. However as the factors of neurosis and extraversion are independent, cheating could be a function of extraversion only, neurosis notwithstanding.

That cheating is a function of hysteria rather than neurosis is supported by Hildebrand's thesis[7] and by Eysenck[4, p. 89] showing that psychopathy is more akin to hysteria than to any other form of neurosis. However, both studies also show psychopaths to be highly neurotic. As against this, Mowrer[9] specifically distinguishes between the neurotic and the psychopath. According to him whereas the psychopath never acquires sufficient super-ego to repress deceitful tendencies, the genesis of neurosis occurs in three stages. First the young child is said to discover that punishment may be avoided through deceit although at a cost of feelings of guilt. Thus, "by the time most children are of school age, they have been powerfully conditioned on this score . . . so that when a child . . . *cheats*, he experiences pangs of conscience"[9]. Second, the pangs of conscience are repressed and, third, "the repudiated sense of responsibility and self-criticism begins to return . . . as 'symptoms' "[9]. Clearly, at least at the descriptive level, Mowrer is not talking about hysteria but about the anxiety states. However it does seem that whereas Eysenck distinguishes between hysteria and the anxiety states from the beginning, Mowrer is postulating a continuum from normality through irresponsibility to anxiety neurosis. If we may regard irresponsibility as part of the hysterical syndrome, as indeed would be apparent from Eysenck's[5] theorizing, then the apparent contradiction between Eysenck's interpretation of Mowrer's account of neurosis and Mowrer's own description is partly resolved. Despite this we are still left with the problem of whether the schoolchild who cheats is simply extraverted, a potential neurotic, a potential hysteric, or none of these. It is the purpose of the present study to examine this point.

METHOD

A cheating test based upon the self scoring technique of Hartshorne and May[6] was administered to 60 school children between the ages of 11 and 13 years. The test consisted of a 24 item arithemetic test which was administered one day, scored over-

[1]This study was made possible by a grant from the Rockefeller Brothers Fund to the American University of Beirut. The author wishes to express his appreciation to the Fund and to Mr. Krikor Maldjian who constructed and administered the arithmetic test.

night, and returned for self scoring the following day on the pretext that the examiner was too busy to carry out the scoring himself. An ordinary school task was used on the assumption that S's would be more likely to falsify answers on such a task in order to escape punishment in the sense proposed by Mowrer.

Estimates of neurosis and extraversion were obtained using a 20 item questionnaire developed from the factor analytic study of Cattell and Gruen [2]. Their factors 2 and 13 were identified as neurosis and extraversion respectively.[2] Although the questionnaire approach does less than justice to Eysenck's viewpoint, similar factors to these have been identified in the realm of objective tests by Cattell [1]. The questionnaire, which contained 10 items to measure each factor, was administered one week after the arithemetic test.

Results

For the total group of 60 subjects the mean number of errors, or opportunities for cheating, was 18.9. Since there were only 24 items altogether it is reasonable to assume that motivation to cheat was quite high. The mean number of alterations was 6.95 although the distribution was markedly skewed with 29 S's making no more than 3 alterations and the remainder spread fairly evenly up to 6 S's making 19 or more illicit corrections. For the purpose of the present study 3 alterations were allowed before the subject was classified as a cheater. In this fashion 31 S's cheated and 29 did not.

Individuals were placed in the neurotic group if they scored 6 or more on the neurotic scale, and in the extraverted group if they scored 5 or more on that scale. This gave N's of 33 and 22 in the two groups respectively. However, although the correlation between the neuroticism and extraversion scales amounted to no more than .21, all but 2 of the extravert group had scores of 5 or more on the neuroticism scale. Thus it was deemed wiser to make comparisons only between the high and low neurotic groups and the high and low neurotic extraverts respectively.

Table 1. Relationship Between Cheating and Personality Questionnaire Data

Factors	Cheaters	Non-cheaters	chi-square
High neuroticism	19	14	1.03
Low neuroticism	12	15	
High neuroticism and extraversion	15	5	6.54*
Low neuroticism and extraversion	16	24	

*Borderline significant at the .01 level

The relevant data in table 1 clearly show no difference in the number of cheaters in the high and low neurotic groups but that a significantly greater proportion of S's scoring high on neuroticism and on extraversion than those with low scores had cheated. Although there were not significantly more cheaters in the neurotic than in the non-neurotic group, those neurotics who did cheat made significantly more alterations than the non-neurotic cheaters. Thus the groups averaged 7.91 and 5.78 alterations respectively which, with sigmas of 7.57 and 4.45, differ significantly at the .01 level. Similar means of 9.25 and 5.62 and sigmas of 6.9 and 7.19 of the neurotic extraverts against the rest show the former to make significantly more alterations at the .001 level of confidence. Clearly, then, neuroticism alone in the children of this group, although related to amount of cheating is unrelated to the number of cheaters, whereas neuroticism and extraversion results in both more cheaters and more alterations by those cheaters. The mean number of opportunities for cheating

[2]Factor 2 was chosen as neuroticism rather than factor 3 as the former seems to fit better with both Eysenck's and Mowrer's descriptions of neuroticism.

in the high neurotic, low neurotic, high neurotic and extravert, and low neurotic extravert groups, were 19.12, 18.74, 19.50 and 18.71 respectively and cannot be held to account for the observed differences. Thus it does seem that extraversion relates more to cheating than does neuroticism.

In order to obtain a further check on the relative effects of neuroticism and extraversion on cheating the high neurotic group was subdivided into high and low extraversion sub-groups respectively. These groups contained 15 and 18 S's of whom 11 and 8 cheated respectively. This distribution although in the direction of showing extraversion rather than neuroticism to be related to cheating yielded a non-significant chi-square when Yates correction was applied. These groups had insignificantly different means of 20.06 and 18.33 errors respectively. When the number of alterations of the two groups was compared, the means of the high and low extraversion sub-groups were 9.13 and 6.89 respectively. Once again although these means are in the direction of showing extraversion to be more important than neurosis with respect to cheating, the difference does not attain statistical significance.

All told, then, our data point to a closer affinity between extraversion and cheating than between neurosis and cheating although whether extraversion alone or in conjunction with neurosis, i.e. hysteria, is the more important cannot be decided from the present evidence.

Summary

Mowrer[9] in presenting a theory of neurosis in opposition to that of Freud has pointed to the role of deception in the genesis of neurosis. However Eysenck[5] has argued that Mowrer and Freud instead of presenting opposing theories of neurosis are really talking about *extraversion* and *introversion* respectively. Thus Eysenck restricted Freud's description of neurosis to the anxiety states and Mowrer's to hysteria. Both this restriction of Mowrer's view and Eysenck's own account of the genesis of hysteria lead to the prediction that cheating in a group of school children would be more related either to extraversion alone or to extraversion and neuroticism than to neuroticism alone. Although more cheating occurred among those S's scoring high on neuroticism, the results were mainly in favor of Eysenck's view. However, it was impossible to tell from the available data if extraversion alone is related to cheating for nearly all S's obtaining high extraversion scores also scored high on neuroticism.

References

1. Cattell, R. B. Personality structures as learning and motivation patterns—a theme for the integration of methodologies. In *Learning theory, personality theory, and clinical research. The Kentucky Symposium.* New York: Wiley, 1954.
2. Cattell, R. B. and Gruen, W. Primary personality factors in the questionnaire medium for children eleven to fourteen years old. *Educ. Psychol. Measmt.*, 1954, 14, 50-76.
3. Eysenck, H. J. *Dimensions of personality.* London: Kegan Paul, 1947.
4. Eysenck, H. J. *The scientific study of personality.* London: Routledge and Kegan Paul, 1952.
5. Eysenck, H. J. A dynamic theory of anxiety and hysteria. *J. ment. sci.*, 1955, 101, 28-51.
6. Hartshorne, H., and May, M. A. *Studies in deceit.* New York: Macmillan, 1928.
7. Hildebrand, H. P. A factorial study of introversion-extraversion by means of objective tests. Unpublished doctor's thesis, London Univers., 1953.
8. Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.
9. Mowrer, O. H. Neurosis: a disorder of conditioning or problem solving? *Ann. New York Acad. sci.*, 1953, 56, 273-288.

A TRANSPOSED FACTOR ANALYSIS OF A GROUP OF SCHIZOPHRENIC PATIENTS[1]

WILSON H. GUERTIN AND RICHARD L. JENKINS

Spring Grove State Hospital
Catonsville, Maryland

Chief, Psychiatric Research
Veterans Administration Central Office
Washington, D. C.

PROCEDURE

The present study is an effort through transposed factor analysis to examine the resemblances among a group of hospitalized schizophrenic patients for any light it may shed upon the nature and classification of schizophrenia. Twenty-nine veterans with schizophrenic psychosis were rated by one of us (W. G.) on the Multidimensional Scale for Rating Psychiatric Patients, Hospital Form (1). These veterans were a series admitted to a Veterans Administration hospital during February to May, 1954. The series included both first admissions and readmissions, and was essentially unselected except for the diagnosis of schizophrenia and accessibility to interview. Mute patients were excluded. In 10 cases, the hospital diagnosis was schizophrenic reaction, paranoid type, in 4, catatonic type, and in 14, chronic undifferentiated type. A normal profile as defined by the pattern of normality of the Multidimensional Scale was also included to make a total of 30 profiles.

Scales 14, 22a, 24, 26, 27, 27a, 44, 51 and 59 were not used in this analysis either because satisfactory data were not obtainable on all cases on these scales or because no useful scatter was found on them in this small series. Cutting points were established on the remaining 55 scales as near as possible to the middle of the distributions. Using the dichotomies so established, a four-fold table was prepared for each of the 435 possible pairings of profiles. The resemblance of each profile to each other profile was then expressed by a tetrachoric correlation coefficient, and a 30 x 30 correlation matrix was prepared. This matrix was factored by the multiple-group centroid method and four factors were extracted. The residuals were negligible and the estimated communality accounted for was well over 90 percent of the variance. These factors were rotated to oblique simple structure and the result is indicated in Table 1. Decimal points have been omitted for convenience. The factors resulting from this transposed analysis have been labelled tS, tA, tB and tE.

RESULTS

In this material, the first and overshadowing factor is tS. Factor tS was reversed in order to have the positive loadings represent pathology. With this reversal there are only two cases with a large positive loading with factor tS. The case with the highest loading on tS shows the most extreme pathology in the entire series, as reflected in the fact this case shows the highest total morbidity scores[2]. Both of these cases carry the hospital diagnosis of schizophrenia, catatonic type. Examination of the rating reveals they both show some disorientation in all spheres, bizarre postures, manneristic movements, and tension. They show moderate perplexity. They speak little in a low or inaudible voice and there is much blocking when they do speak. They are somewhat troubled by vague fears. They also show some unwarranted feelings of sinfulness and are self-depreciatory. On the ward they do not eat well (one requires spoon feeding), they are underactive, movements are slowed and they seem tired. They require help in the daily routine of dressing, going to the toilet, etc. Their verbal responses are very limited. They are sloppy and dirty. They isolate themselves, do not speak to others and show no interest in any ward activities.

---

[1]From the Veterans Administration Hospital, Knoxville, Iowa, and the Veterans Administration Central Office. The authors are indebted to Dr. Maurice Lorr for his statistical advice.

[2]The total morbidity score is a total of the deviations on the Multi-dimensional Scale for Rating Psychiatric Patients of deviations from the normal pattern with each deviation weighed by the number of such intervals it deviates.

TABLE 1. CORRELATIONS OF 29 SCHIZOPHRENIC CASES AND ONE HYPOTHETICAL NORMAL CASE WITH THE ROTATED PERSON FACTORS. DECIMAL POINTS ARE OMITTED.

Case No	Diagnostic Classification	Factors tS	tA	tB	tE	Total Morbidity Score
29	Catatonic	79	27	29	26	67
9	Catatonic	59	47	26	06	50
28	Paranoid	-54	06	38	15	23
23	Chronic un.	-54	48	-02	00	33
18	Chronic un.	-56	08	43	03	24
5	Paranoid	-60	06	35	-13	40
15	Paranoid	-64	31	-09	50	28
17	Paranoid	-64	03	12	23	25
21	Chronic un.	-65	36	-03	05	29
14	Chronic un.	-69	19	36	07	15
20	Paranoid	-75	28	-01	02	23
25	Chronic un.	-86	35	04	01	28
24	Chronic un.	-90	00	27	30	17
30	Normal	-93	17	-06	25	0
26	Catatonic	02	62	-02	-07	53
10	Chronic un.	39	49	41	32	18
4	Chronic un.	40	42	41	34	49
8	Chronic un.	-02	-37	93	06	35
12	Chronic un.	15	06	82	31	29
19	Catatonic	45	16	63	-16	25
7	Chronic un.	28	01	30	14	30
13	Paranoid	34	06	43	14	35
16	Catatonic	19	09	31	19	25
27	Chronic un.	-39	-07	40	35	43
6	Paranoid	-02	31	23	90	34
11	Paranoid	-03	-03	02	89	52
2	Paranoid	34	-02	27	70	37
1	Paranoid	25	14	32	67	37
22	Chronic un.	-08	04	39	57	38
3	Chronic un.	33	08	50	54	28

Factor tS, at its negative end, is clearly a factor of normality. The profile loaded most negatively with factor tS is the pattern of normality. The three patients with the next highest negative loading with this factor have identical ratings on 17 interview scales and 14 ward behavior scales. On all but three scales these identical entries are in perfect agreement with the pattern of normality. These differences were that these three patients were considered to have less in the way of mood swings than the average, to under-react in emotional response and never cursed or used profanity. Although each of these patients is schizophrenic, they did not *as a group* differ from the pattern of normality in any other regard except in showing less than normal interest in ward activities. Factor tS, at its negative end, accounts for the resemblances between our patients which are due to their individual resemblances to normal individuals.

Factor tS seems to be a factor representing chiefly the *degree* of schizophrenic pathology, or, viewed in more general terms, this is a common species factor expressing the central tendencies of the group of subjects. The two cases highest on this factor show both disorganization and withdrawal and the more extreme shows belligerence as well.

The remaining three factors do not relate substantially to the severity of the schizophrenic disorder in that these factor scores do not correlate significantly with the morbidity score, but relate rather to the kind or quality of the disorder shown. If we examine the ratings of the patients in our series with the highest loadings on factor tA, we find these patients are usually conspicuously tense, and tend to be

under-talkative. There is some blocking when they do speak. They show manneristic movements and grimaces. There is very little body concern and no use of complaints. While they resemble those patients with the highest loadings on tS in a number of regards, their speech is more relevant, and they do not give evidence of feelings of sinfulness, nor are they self-depreciatory. They present no special problem in ward care and are not assaultive, destructive or belligerent. The patient highest on factor tA was diagnosed catatonic. Factor tA appears to be a factor of withdrawal without characteristic gross disorganization. It may well be that this factor would be better defined and better represented had it not been necessary to exclude mute patients from the study.

Patients rating high on factor tB show no perplexity or vague fears. They are not tense. They talk but little. Disorientation is frequent, particularly disorientation for people. There are no feelings of sinfulness. Reactions are usually a little slow. Hostility tends to be low. These patients are not recognizably anxious, although there may be some use of complaints. Some have auditory hallucinations, but they are little given to morbid suspicion or to ideas of reference or influence. Emotional response tends to be inadequate. On the ward they seem worn out but some are occasionally irritable, uncooperative or resistive. They take some interest in their surroundings. The two patients highest on this factor were diagnosed schizophrenic reaction, chronic undifferentiated type.

If we turn now to factor tE, we find that these patients are characterized by a prominence of perplexity and/or morbid fears which tend to disrupt their activities. Speech shows some blocking but they are as talkative as the average and there is no stereotyped speech. Anxiety is usually evident. They show a pronounced tendency to be preoccupied with their health or the functioning of their body organs and they may use complaints to get attention. They are usually self-depreciatory and may show unwarranted feelings of sinfulness. There are no bizarre postures or manneristic movements. Their feeling-response is adequate. On the ward they eat poorly and have trouble sleeping. They take interest in ward activities and speak to others at least occasionally. They are not slow in movement and do not appear tired. Some are occasionally resistive or assaultive or destructive. They are distinguished chiefly by various evidences of anxiety or agitation. The four patients classifying highest on this factor are all classified as paranoid, but there are paranoid patients not high on this factor and there are patients high on this factor who are not paranoid.

The correlations between the oblique factors are .16 for tS with tA, −.15 for tS with tB, .05 for tS with tE, −.45 for tA with tB, .04 for tA with tE, and .05 for tB with tE. It will be noted that the only substantial correlation is the negative one between tA and tB.

In Table 2, the cases highest on each of the four factors of this transposed analysis are described in terms of 11 factors resulting from a direct analysis of the scales and presented by Lorr [1]. Since there were only two cases with large positive loadings on factor tS, these two cases are characterized in Table 2. In the cases of the factor tA, the characterization of the three cases with the highest loadings with this factor was entered, and the same is true for factors tB and tE.

The cases with the highest loadings of factor tS show severe degrees of pathology on most of the factors, although paranoid projection is not pronounced and one of these patients is belligerent while the other is submissive. The deviation from the normal is most pronounced on retardation, resistiveness, motor disturbances, withdrawal, self-depreciatory tendencies and conceptual disorganization.

The cases highest on factor tA also show motor disturbances, retardation and reduced activity level, withdrawal and agitation. Although they also show conceptual disorganization and the highest level of perceptual distortion, the "withdrawal" type of symptoms are the most consistent.

The cases highest on tB show less of the factor of melancholy agitation than does the average normal. They are low on paranoid projection and motor disturb-

TABLE 2. SCALE (MSRPP) CHARACTERISTICS OF THE PATIENTS BEST DEFINING EACH OF THE FIVE PERSON FACTORS

Scale Factors	Person Factors tS	tA	tB	tE
A. Manic Excitement *vs.* Depressive Retardation	Retarded	Retarded	Retarded	Slightly manic
B. Resistiveness	Resistive	————	Slightly Resistive	Slightly resistive
C. Paranoid Projection	Slightly Paranoid	Paranoid	————	Paranoid
D. Activity Level	Inactive	Inactive	Inactive	Slightly overactive
E. Melancholy Agitation *vs.* Apathy	Agitated	Agitated	Slightly apathetic	Highly Agitated
F. Perceptual Distortion	Hallucinatory	Hallucinatory	Hallucinatory	Hallucinatory
G. Motor Disturbances	Highly manneristic	Highly manneristic	————	Slightly manneristic
H. Belligerence *vs.* Submissiveness	Variable	Submissive	Submissive	Variable
I. Withdrawn *vs.* Social	Severely withdrawn	Withdrawn	Withdrawn	Slightly withdrawn
J. Self-Depreciatory *vs.* Grandiose	Self-depreciatory	————	————	Self-depreciatory
K. Conceptually Disorganized	Disorganized	Disorganized	Disorganized	Disorganized

Blank spaces indicate no characteristic difference from normal.

ances, are retarded, inactive, and unbelligerent, show no self-depreciatory tendencies, but do show perceptual distortion and conceptual disorganization.

The cases highest on tE are not retarded as are those high on all other factors but, on the contrary, show slight manic tendencies, increased activity level, the highest level of melancholy agitation, some self-depreciation and are the least withdrawn.

DISCUSSION

Previous work published on factor analysis of these scales has been a product of direct analysis, rather than of transposed analysis, and the results of the two approaches are, of course, not directly comparable. Direct analysis of the inter-correlations of *scales* usually produces a larger number of first-order factors than does a transposed analysis, which deals with the inter-correlations of *persons*. However, a second-order analysis of the inter-relations of the oblique factors obtained from a direct first-order analysis produces second-order factors which may compare in number with those obtained by a transposed analysis, and which seem to bear a very interesting relation to these factors. Degan's analysis of Moore's material indeed strongly suggests that the second-order factors may relate themselves to clinical diagnoses.

Factor tS is clearly a bipolar factor ranging from normality at the negative end to an extreme of schizophrenic disorganization at the positive end. The remaining three factors relate rather to varieties of schizophrenic pathology than to degree of pathology.

Factor tA seems to be chiefly a factor of schizophrenic withdrawal, although patients who isolated themselves so effectively they could not be interviewed were excluded from the study. Factor tA seems to be closely related to second-order factor Z obtained by Degan[2] in his direct factor analysis of the material of Dr. Thomas Verner Moore, to second-order factor Alpha of Lorr, Jenkins and Holsopple[3], to second-order factor AA of Lorr, Jenkins and O'Connor[4], and to second-order factor A of Lorr, Holsopple, Jenkins and O'Connor[5]. Degan related his factor Z to catatonic schizophrenia while Lorr and his co-workers have described their factors as resistive withdrawal with motor disturbances.

Factor tB appears to relate to a schizophrenic disorganization without tension, a disintegrating process without apparent force or heat, which is presumably either dependent upon a poor personality organization or which marks a late stage in a process which has largely burned itself out. Factor tB seems to mark a result more than a process. Factor tB seems to bear a close relation to second-order factor X of Degan's analysis[2], which he related to hebephrenic schizophrenia, to second-order factor Beta of Lorr, Jenkins and Holsopple[3], which was considered to be a factor of perceptual and thinking distortion, to second-order factor BB of Lorr, Jenkins and O'Connor[4], considered a factor of schizophrenic disorganization, projection and distortion of thinking and perception, and to second-order factor C of Lorr, Holsopple, Jenkins and O'Connor[5], which was considered a factor of thinking disorganization and perceptual distortion.

Factor tE appears to be a factor of schizophrenic agitation and anxiety. It is clearly closely related to factor B of the lobotomy study of Lorr, Jenkins, Holsopple and O'Connor[5], which was considered to represent schizophrenic excitement. It seems also to bear some relation to the agitated end of the bipolar factor CC of Lorr, Jenkins and O'Connor[4], which ranges from apathetic withdrawal to agitated hyperactivity.

SUMMARY

An inverse factor analysis of the ratings of 29 schizophrenic patients on the Multidimensional Scale for Rating Psychiatric Patients revealed four factors. These were rotated to oblique simple structure.

One factor appears to be a general schizophrenic factor relating to the degree of schizophrenic pathology. Three other factors relate to a tense schizophrenic withdrawal, to a schizophrenic disintegration without tension, and a schizophrenic agitation and anxiety.

REFERENCES

1. *Veterans Administration Technical Bulletin 10-507.* Multidimensional Scale for Rating Psychiatric Patients, Hospital Form (By Dr. Maurice Lorr).
2. DEGAN, J. W. Dimensions of Functional Psychosis. *Psychometric Monograph*, No. 6, 1952.
3. LORR, M., JENKINS, R. L. and HOLSOPPLE, J. Q. Factors Descriptive of Chronic Schizophrenics Selected for the Operation of Prefrontal Lobotomy. *J. Consult. Psychol.*, 1954, 18, 293-296.
4. LORR, M., JENKINS, R. L. and O'CONNOR, P. Factors Descriptive of Psychopathology and Behavior of Hospitalized Psychotics. *J. Abnorm. Soc. Psychol.*, 1955, 50, 78-86.
5. LORR, M., HOLSOPPLE, J. Q., JENKINS, R. L. and O'CONNOR, P. Factors of Change in Lobotomized Chronic Schizophrenic Patients. *J. Consult. Psychol.*, 1955, 19, 39-43.

AN IMPORTANT NEW MONOGRAPH

New Approaches to Psychotherapy Techniques
by
ALBERT ELLIS

A comprehensive interpretive review of the periodical literature (1950-1953) covering all important contributions to the field. Reprinted from this Journal, a July 1953.

PRICE: $2.00 ORDER NOW

JOURNAL OF CLINICAL PSYCHOLOGY
5 Pearl Street Brandon, Vermont

MEASURES OF INTELLIGENCE IN BRAIN DAMAGED CHILDREN

JAMES J. GALLAGHER E. PAUL BENOIT
AND
HERBERT F. BOYD
University of Illinois

PROBLEM

Workers in the field of exceptional children have long been seeking measuring instruments adapted to the needs and functioning abilities of exceptional children. All too often, standard instruments such as the Stanford Binet cannot provide correct or accurate information on the exceptional child's potential for learning. In this connection, the recent release of the Columbia Mental Maturity Scale[1] for clinical use with children has been a promising event. Because of its ease of administration and its applicability to practically all types of handicapped children it stands a good chance of becoming a popular and much used instrument.

The purpose of the present study was to determine if the CMMS compared favorably with the Stanford Binet and the Leiter International Performance Scale[2] when applied to an atypical population of brain damaged, mentally retarded children. This problem is quite pertinent since the CMMS was validated apparently only by means of a comparison with the Binet on an incompletely described school population. It is the writers' contention that there is room for doubt that relationships derived from such a sample will hold true for a deviant group. Since the CMMS is supposed to apply to the brain damaged children, it would be worthwhile to investigate its actual relationship to the other tests on such a sample.

Previous studies. With the exception of the Binet and Wechsler tests, there is a surprising paucity of information concerning the standardization of the different individual tests of mental ability. The only report available on the CMMS so far consists of information available in the manual (5). The standardization population was composed of 957 'presumptively normal' children with a CA range of from three to twelve years. The correlations with the Binet for all age groups were reported as .78 but the standard deviations for the CMMS at the lower age levels were reported as being somewhat greater than those of the Binet.

No data seem to have been published as yet on the original standardization of the LIPS. The Caucasian norms for the LIPS (4) were established by comparing the results on this performance scale with IQ standing as shown by a battery of previously administered group tests. The Arthur adaptation of the LIPS (1) relates its results to those of the Arthur Point Scale; but such a comparison has little meaning, as we have scant knowledge about the original sample that Arthur used in developing her test.

Bessent (2) obtained a correlation of .92 between IQ scores on the Binet and LIPS but this figure is of doubtful value, owing to the wide CA range (4 to 58), small sample (N = 20), and the atypical nature of the sample (50% of Mexican ancestry).

METHOD

Subjects. The forty children in this experiment were drawn from the residential population of the Dixon State School at Dixon, Illinois. There were twenty-six boys and fourteen girls, ranging in CA from 7-4 to 13-10. The diagnosis of brain injury was made by means of the Riggs and Rain classification system (6), which consists of a systematic evaluation of case history material. Any child having parents or siblings who were retarded were discarded from the population. As the present data were gathered as part of a long range training experiment, no cases of marked visual or auditory difficulties or severe emotional disturbance were included.

---

[2]Hereafter referred to as LIPS
[1]Hereafter referred to as CMMS

Procedure. The three tests were administered as part of a comprehensive battery designed to evaluate the children before the start of their educational program. The children were brought to a central building at the institution, and were given the tests in a more or less random order. All of the children had come to the building before in order to participate in group play periods and hence readily adjusted to the situation. Rapport was easily established with the great majority and lack of cooperation was not considered a serious problem by the examiners. Standard t tests were calculated on the differences between correlated means described by Guilford[3]. The Pearson product-moment correlations were corrected for restriction of range in the present population.

TABLE 1. COMPARISON OF IQ SCORES OF INSTITUTIONALIZED BRAIN DAMAGED CHILDREN ON THREE MEASURES OF INTELLIGENCE

Test	N	Mean IQ	σ	Range of Mental Age
Stanford Binet (L)	40	48.85	7.36	2-9 to 6-10
Leiter†	40	45.50	11.41	2-3 to 7-0
Columbia*	40	57.34	15.09	3-1 to 12-0

*Difference between means of Columbia and other measures significant at 1% level of confidence.
†Difference between means of Leiter and Binet significant at 5% level of confidence.

RESULTS

The correlations obtained on this population were as follows: Binet *vs.* CMMS = .93, Binet *vs.* LIPS = .86, CMMS *vs.* LIPS = .85. While this compares favorably with past studies, correlation is only one criterion for measuring comparability of tests. Table 1 reveals a comparison of the mean scores of the present group. There was an average mean difference of nine IQ points in favor of the CMMS over the Binet and this difference was statistically significant at the one per cent level of confidence. The LIPS mean fell below the Binet mean by somewhat more than three IQ points, a difference significant at the five per cent level of confidence. Thus the means of all three measures differ at a statistically significant level from each other. It is also interesting to note that the CMMS yielded a much wider range of MA scores than did either of the other two tests.

TABLE 2. DISCREPANCY OF IQ SCORES OF NON-VERBAL TESTS WITH STANFORD BINET (L)

Test	N	No. of cases showing discrepancies with Binet in terms of IQ Scores						
		Below −15	−15 to −11	−10 to −6	−5 to +5	+6 to +10	+11 to +15	+15 & above
Leiter	40	5	5	6	19	2	3	
Columbia	40	1	2	4	6	10	6	11

Table 2 displays in more concrete fashion the trend of discrepancies between the Binet and the other two tests. Although the means of the Binet and LIPS could be considered statistically different, only twelve per cent of the cases deviated as much as fifteen IQ points from the Binet, and almost fifty per cent of the LIPS scores were within plus or minus five IQ points of the Binet. In contrast to this situation, thirty per cent of the CMMS scores deviated from the Binet by fifteen or more IQ points and only fifteen percent of the cases were within plus or minus five IQ points of the Binet. There seems to be little doubt that there is a distinct lack of comparability between the Binet and the CMMS on the present population.

A further question was raised; namely, does the general mental level of the child have any influence on the amount of inter-test discrepancies? Table 3 reveals the relationship between discrepancies and mental level. The CMMS appears to agree fairly well with the Binet when the child has a Binet MA of less than 4-7. However, when the mental level is above that point, the discrepancies become quite marked until they reach an astounding average difference of nineteen months of mental age above the five and a half year level. In contrast to these findings, the LIPS seems to show its best agreement with the Binet above the four and a half year level and only becomes markedly discrepant with the Binet when the child is below that mental level.

TABLE 3. RELATIONSHIP OF MENTAL LEVEL TO DISCREPANCIES IN MENTAL AGE SCORES BETWEEN STANFORD BINET (L) AND NON-VERBAL SCALES

Tests	Binet MA 4–7 & below	Binet MA 4–8 to 5-5	Binet MA 5–6 & above
	N=12	N=16	N=12
Leiter	-7.7	-3.2	-2.5
Columbia	+3.7	+10.9	+19.2

DISCUSSION

The present results strongly suggest that the Binet, CMMS, and LIPS are not comparable on the present population. This is particularly true of the LIPS when used on children below a Binet mental level of four and a half, and on the CMMS above that level. While this finding does not rule out the use of any of these instruments, it does emphasize the necessity of validating them independently of each other.

It is difficult to understand why so much time is invested in developing a scale such as the CMMS only to have the crucial problem of validity shrugged off with a routine comparison with the Binet. Certainly the most important aspect of a test for exceptional children is that it validly measures or predicts some sort of vocational, social, or educational success or adjustment. These criteria have not been applied with sufficient rigor to the present non-verbal scales.

Two major criticisms of the CMMS apart from this study may be mentioned. First, their standardization population seems to be selected exclusively from New York State, if not New York City. Second, the CMMS IQ scores are not directly comparable to the Binet (the authors recognize this) since the standard deviations for the CMMS average about four IQ points greater than the Binet. All other things being equal, that would mean that we should have expected higher scores on the Binet than the CMMS for our group. That is, a position of –2 sigma on the normal distribution would be represented by a Binet IQ of 68 but, theoretically, a CMMS IQ of 60. Our results are directly contrary to this supposition and should make the practicing clinician cautious in using these non-verbal scales to the exclusion of other measures.

Is there any educational significance to cases with high intertest discrepancies? Do the scales predict educational or training aptitude? Are these results merely peculiar to our sample? These remain unanswered questions but it is hoped that some light will be shed on them when the total training experiment is completed.

SUMMARY

Forty institutionalized and mentally defective brain injured children with a CA range of 7 to 14 were given the Stanford Binet (L), the CMMS, and the LIPS. Results indicate the need for extreme caution in assuming comparability of test scores of these three measures on our present population. Some possible explanations for these discrepancies are briefly discussed.

References

1. Arthur, Grace. The Arthur adaptation of the Leiter International Performance Scale. *J. clin. Psychol.*, 1949, *5*, 345-349.
2. Bessent, T. E. A note on the validity of the Leiter International Performance Scale. *J. consult. Psychol.*, 1950, *14*, 234.
3. Guilford, J. P. *Fundamental Statistics in Psychology and Education.* New York: McGraw-Hill, 1950.
4. Leiter, R. G., Caucasian norms for the Leiter International Performance Scale. *Psychol. Serv. Cent. Journal*, 1949, *1*, 136-138.
5. Lorge, I., Blum, Lucille, and Burgemeister, Bessie. *Manual for the Columbia Mental Maturity Scale.* Yonkers-on-Hudson; World Book, 1954.
6. Riggs, Margaret and Rain, Margaret. A classification system for the mentally retarded. Part I: Description. *Training Sch. Bull.*, 1952, *49*, 75-84.

A STATISTICAL STUDY OF FREUD'S THEORY OF PERSONALITY TYPES

ROSS STAGNER AND J. WELDON MOFFITT

University of Illinois *Brigham Young University*

Problem

One of the chronic problems in personality theory is that of the best way to conceptualize personality organization. Trait theorists have generally held that organization can be conceived in terms of generalized ways of perceiving and responding to a class of situations. The organizing principle is derived from the learning process, so that functionally related habits become organized into higher-order habits. Different traits are relatively independent of one another. Type theorists, on the other hand, have argued for an over-all unity of personality, a Gestalt quality, such that all the parts have a certain necessary relationship to one another. The organizing principle may be derived either from learning or from heredity, but in either case it is assumed to cover a much more inclusive personality configuration than a specific trait.

Freud proposed the classification of certain individuals as personality types on the basis of the relative dominance of anal fixations. He wrote: "Among those whom one tries to help by means of psychoanalytic treatment, one very often meets with a *type of character* in which certain traits are very strongly marked while at the same time one's attention is arrested by the behavior of these persons in regard to a certain bodily function . . . " (1, p. 45, italics ours.) Freud later noted the existence of an "oral type" and this concept was elaborated by Abraham (2). Throughout the psychoanalytic literature we find phallic types, narcissistic types, early and late oral types, anal-retentive and anal-expulsive types, and so on.[1] It seems a legitimate deduction that, if a "type" is a functional, organized unity, then individuals classified into this type should show greater similarities to one another than to persons classified into other types. The technique of correlating persons, therefore, should be a suitable device for testing the appropriateness of the type-theory.

An instrument which is admirably adapted to testing the Freudian type-theory is Krout's Personal Preference Scale (KPPS). (3). The authors differentiate ten "developmental stages" in personality growth, something of a refinement of orthodox theory, and provide ten test items presumed to be diagnostic of important influences

[1]It is especially interesting to note that in his 1931 paper on "libidinal types" (*Collected Works*, Vol. V, 247-251), Freud used a totally different basis for postulating types, and completely ignored his earlier writings within the framework here described.

derived from each of these stages: infantile passive ("prenatal"), oral sucking, oral sadistic, anal retentive, anal expulsive, narcissistic, feminine, masculine, intrafamilial sublimation and social sublimation. Factor analyses of the intercorrelations of representative items from the ten subtests by Stagner, Lawson and Moffitt [4] give encouraging although not entirely clear-cut support to the hypothesis that ten unified trends in behavior are present and that they correspond roughly to the patterns theoretically derived by Krout and Tabin. Since the 100 items utilized by Krout and Tabin relate to the different "developmental stages" of Freudian theory, they should offer a suitable medium for testing the typological conception. Our statistical test is this: Persons selected as representatives of one type should correlate more highly among themselves than with persons selected as representatives of other types.

Procedure

We first administered KPPS to 200 male students at the University of Illinois. These were scored and the three men who were lowest (most typical) on each subtest were selected as examples for the corresponding "type". KPPS has 100 items which were answered on a 5-step scale.[2] Pearson correlations were therefore computed between each S and every other S, giving us a matrix of 435 coefficients. Inspection indicated some support for the prediction that the correlations within a "type" were higher than those between types.

To test the significance of the difference we made use of an analysis of variance procedure suggested by Dr. L. J. Cronbach. This involved getting estimates of within-groups and between-groups variance. Since it would be improper to include the person's correlation with himself, we included as within-groups data the person's correlations with his two fellow-examples. These were transformed into Fisher's z, as were the other coefficients used. To meet the requirement of independent data in the between-groups sum of squares, we included in our analysis only the correlation of one pair of subjects for each pair of types. To clarify: subject 1 was correlated with 4, 7, 10 . . . 28; subject 2 was correlated with 5, 8, 11 . . . 29; and subject 3 with 6, 9, 12 . . . 30. Thus, for the relationship of "type A" to "type B", the three coefficients used would be 1-4, 2-5 and 3-6. For any given F-ratio, therefore, we would have 6 sums of squares for within-variance and 3 sums of squares for between-variance. For within-groups variance there were 23 degrees of freedom, for between-groups, 5. This was further broken down into 1 for rows, 2 for columns and 2 for interaction.

To test the prediction of significantly higher within-type correlation,[3] we started with those types showing the highest mean intercorrelation within the type and the largest difference from the mean correlation with the 9 other types. The four highest types were tested. In every instance the F-ratios were non-significant, the highest F obtained being 2.03. This is well below the figure required for the 5% level of significance. Since these were the four most favorable cases, the others were not analyzed.

Discussion

The data are interpreted as rejecting the hypothesis that a typological organization of personality is the best way of representing the functional relationships of these particular responses. In a broader way, they cast additional doubt upon the general Freudian typology, and in a still broader frame of reference, they bear on the general controversy as regards the legitimacy of type-theories of any kind. Obviously such extensions of the interpretation are very tentative.

It may be argued by defenders of Freudian typology that our data do not support the type-theory but that they are inadequate for a rejection of the hypothesis of psychosexual types. The typologist may hold that only relatively rare cases (prob-

[2] In the cuurent edition of the test, only a 3-step answer scale is utilized. The present analysis is based on cases from the same population reported by Stagner, Lawson and Moffitt [4].

[3] More precisely, the test applies to the null hypothesis, that the mean correlations do not differ significantly. Our findings then indicate that the null hypothesis cannot be rejected.

ably neurotics) exemplify the type. To this we would reply that such rare instances may be only those special combinations of varying trends which must be expected, in accordance with probability theory, to occur occasionally in any population. We do not concede the appropriateness of using them to justify a theory of personality organization. Another objection may be that the items on KPPS are not adequate to define the types postulated. To this Krout and Tabin [3] provide tentative answers in terms of the clinical derivation of the items. Such an objection, of course, can never be answered conclusively.

Our interpretation is that the so-called oral erotic, phallic and other "types" are to be considered only as extremes of a trait continuum. An individual may very well experience over-gratification or deprivation at a specific stage of development, which may foster the establishment of fixed ways of perceiving and responding. These relatively persistent patterns become nuclei of organization around which later characteristics develop. However, the degree of generalization involved is quite restricted, and it is completely misleading to make such a trait the basis for a type-classification. The trait may be highly salient in therapy, or even in everyday observation, but for the general population it is not a cardinal principle of personality organization around which other personal manifestations cluster in subsidiary structures. If the observer knows that a given subject is high on oral-eroticism, this knowledge does not provide a firm basis for predictions regarding other important traits of personality.

Summary

1. It was hypothesized that, if Freud's theory of personality types were valid, persons selected as examples of one type should intercorrelate among themselves significantly higher than with those selected as examples of other types.

2. Using Krout's Personal Preference Scale as a measuring instrument, three men were chosen to exemplify each of ten "types."

3. Intercorrelations were computed among all cases. By an analysis of variance technique, a test was made of the hypothesis that the mean within-type correlation did not differ significantly from the mean between-types correlations. The statistical tests indicated no significant differences for any of these comparisons.

4. The data are interpreted as casting doubt on the appropriateness of using typological formulations as proposed by Freud and some of his followers. A concept of independent traits corresponding to the various alleged "types" seemed to be more defensible.

References

1. Freud, Sigmund. Character and anal erotism. In: *Collected Papers*, vol. 2. London: Institute of Psychoanalysis, 1924.
2. Abraham, Karl. The anal character. In: *Selected Papers on Psychoanalysis*. London: Institute of Psychoanalysis, 1927.
3. Krout, Maurice H., and Tabin, Johanna K. Measuring personality in developmental terms. *Genet. psychol. Monogr.*, 1954, 50, 289-335.
4. Stagner, Ross, Lawson, Edwin D., and Moffitt, J. Weldon. The Krout Personal Preference Scale: a factor-analytic study. *J. clin. Psychol.*, 1955, 11, 103-113.

CALL FOR PAPERS

This Journal is now able to offer immediate publication for high priority research reports. Increased efficiency of publication processes now permits publication of acceptable papers in three months from date of receipt. Submit your studies now.

RELIABILITY OF THE DESPERT FABLES, A STORY COMPLETION PROJECTIVE TEST FOR CHILDREN

HELEN E. PEIXOTTO

Catholic University of America

Introduction

The Despert Fables is a projective test of the content type which is particularly appropriate for children. It consists of a number of stories which the child must complete, thus each story ends with a question. The original set of incomplete stories was published by Düs[2] and then was translated into English by the psychiatrist, Louise Despert[1]. The technique was further modified and lengthened by Fine[3], and it is this last revision which has been used in this study.

This test is particularly useful because it can be used with children as young as 3 years of age—possibly younger. It is well known that the child's first response to pictures is description. This is the successful response on the Stanford-Binet at the 3½ year level.[1] Between 3 and 4 years of age description is the developmentally normal type of response to pictures and, from 4 to 6, it may well represent retardation rather than evasion to threatening stimuli. The story technique, therefore, has certain advantages over the more common thematic tests which use pictures for stimuli. In the Fables the child completes a story and thus reflects psychodynamic material and feelings relevant to a situation such as the story portrays.

The present study is an attempt to investigate the reliability of the Fables. Because of the structure of the test, 20 discrete Fables, the only type of reliability which is meaningful is that of consistency, i.e., will the test elicit the same information on successive administrations, all other things being equal.

Procedure

The Fine revision of the Despert Fables was administered as a group test to classes three through eight in an elementary school.[2] The 20 Fables were mimeographed and below each Fable there were three blank lines for the child's response. The instructions were printed at the top of each sheet and were also read to the class. The children were not told at the end of the first administration that they would be retested. The second administration was six weeks after the first. The instructions given were the same as for the previous administration. Questions were answered non-directively.

The subjects were 177 middle class, white, normal children. Although there is a nominal tuition fee, the children are from the neighborhood and represent a fair cross-sample of the population. Table 1 gives the pertinent statistics for the subjects.

Table 1. Description of Subjects

Grades	3	4	5	6	7	8	Total
Girls	14	16	21	16	15	11	93
Boys	14	15	13	12	15	15	84
Totals	28	31	34	28	30	26	177
Age - Range	8-9 yrs.	8-11 yrs.	10-11 yrs.	11-12 yrs.	11-13 yrs.	13-14 yrs.	

[1]On the other hand, a well known type of evasion used by both older children and adults when the stimulus material becomes threatening is that of description.

[2]Campus School run by the Catholic University of America for teacher training and situated in Washington, D. C.

The technique used was the test-retest method for obtaining reliability. Thus the Fables are scored "same" or "different" according to whether the child gave the same or a different response on the retest as he had given on the initial administration. In scoring these data one could demand identity in order to score the responses as "same" or one could be less rigorous and score a response as "same" if the response on the retest represented the same psychodynamics as the response on the first test. For example, a Fable investigating dependency could have several responses which would be categorized as dependent or independent but would not, necessarily, be a verbatim repetition of each other.

Therefore, the data were scored in two ways, first they were scored according to whether or not the response was a verbatim reproduction, and secondly, on those Fables failing to attain satisfactory reliability, whether or not the response represented the same psychodynamics as given on the first administration of the test. In the first instance it is obvious that the criterion for scoring is sufficiently objective, that no subjective element would vitiate the results. The same cannot be said for the second method. There were instances where the particular dynamic formulation of the Fable was in question, others where the dynamics represented by a particular response was subject to interpretation. All Fables and possible responses were presented to judges before the second scoring was undertaken. Five judges participated: two clinical psychologists and three graduate students in clinical psychology. Unanimity was achieved regarding the psychodynamic most obvious for each Fable and the interpretation of responses. All of the judges were familiar with the Fables and were currently using them in a clinical setting.

Results

It will be noted from Table 1 that the age span of the subjects is wide, 8 to 14 years. It is well known that developmental changes in this period from middle childhood to early adolescence are great. This is true in all areas of development: physical, glandular and psychodynamic. Since different Fables assess different psychodynamics it was hypothesized that different Fables might elicit psychodynamic responses reliable at one age but not reliable at another age. Thus the reliability was found for the entire group and also for three age groups, 8-11 yrs; 10-12 yrs; 11-14 yrs; or grades 3 and 4; 5 and 6; 7 and 8. The data for these analyses are presented in Table 2.

Reliability was estimated by the Phi Coefficient. As indicated above, Phi Coefficients were obtained for the group as a whole, and for the grades paired as shown in Table 2. Three things are obvious immediately from Table 2: first, in every instance the number of Fables showing significance of reliability above the 5% level is well above chance; second, the younger children give more reliable test performance than do the older children; and third, there is considerable variability among the Fables themselves. Fables 1, 2, 8 and 14 are either lacking in stability or are of very limited stability, while Fables 9, 11, 15, 16 and 20 can be considered reliable only for the youngest group of children. Fables 5, 12 and 15, on the other hand, are more reliable for the older children than for the younger children. Fable 4 is unique being reliable for the intermediate group, but not for the older or younger group. The other Fables, namely: 3, 6, 10, 13, 17 and 19 show adequate, across the board reliability, that is, seven of the Fables are reliable for all the age groups tested.

As indicated above, it was not necessary to compute reliabilities for the entire group of twenty Fables according to the consistency of the dynamics expressed, but only for those which had not shown satisfactory reliability on verbatim responses. These Coefficients are given in the last column of Table 2, headed Dynamics. Since these results were so disappointing for the entire group, this was not attempted for the individual grades. In fact, little confidence seems to be added to the reliability of any Fable with the possible exception of the second, as a result of this analysis.

An attempt was made to obtain a multiple Phi Coefficient[4] but this was not successful. The resultant ϕ_T was .008 which is obviously insignificant. It was evi-

TABLE 2. SIGNIFICANCE OF RELIABILITIES BASED ON TEST-RETEST OF FABLES

	Identical Responses				Dynamics
Fables	Grades 3-4	Grades 5-6	Grades 7-8	Grades 3-8	Grades 3-8
1	*	*	*	*	.02
2	*	*	*	*	.001
3	.001	.001	.001	.001	
4	*	.001	*	.001	
5	*	.01	.01	.01	
6	.001	.001	.001	.001	
7	.001	.001	*	.001	
8	*	*	*	.05	.10
9	.001	.05	*	.001	
10	.01	.01	.05	.001	
11	.01	*	*	.05	
12	.01	*	.001	*	
13	.001	.001	.01	.001	
14	*	*	.05	*	.01
15	.05	*	.01	*	
16	.01	*	*	.05	
17	.001	.01	.001	.001	
18	.01	.001	.001	.001	
19	.001	.01	.001	.001	
20	.01	*	*	*	
No. Fables Reliable	14/70%	11/55%	11/55%	14/70%	

*Denotes significance less than .05.

dent that varying the combinations to obtain ϕ_T would give different results. Thus it may have been possible to manipulate the data in such a way that eventually one would have a significant coefficient. However, it was thought that reliability under these circumstances would have such little significance that this attempt was not pursued further.

DISCUSSION

Since we have used Fines[3] revision of the Despert Fables we will limit a discussion of the results to his interpretation of the twenty stimuli. According to his findings the test measures eight different psychodynamic variables. These are indicated in Table 3 to facilitate comparison with the previous Table. It is not felt that these variables are beyond dispute, but for the purpose of this study we can accept them as the basis of discussion.

TABLE 3. PSYCHODYNAMICS MEASURED BY THE VARIOUS FABLES (ACCORDING TO R. FINE).

Psychodynamic	Fables
Dependency	1, 11, 14, 17
Hostility	5, 8, 12
Identification	4
Sibling rivalry	6, 20
Rejection	2, 3
Castration fears	10
Oedipus complex	15
Fears and wishes	7, 9, 13, 16, 17, 19

Where there is more than one stimulus associated with a particular variable or dynamic, in no case did we find 100% reliability for this variable. Thus in the case of dependency, Fable 17 is the only one consistently reliable, while Fable 11 is re-

liable for youngest children. In the case of hostility, Fable 5 is reliable for the intermediate and older children while Fable 12 tests this variable reliably for the youngest and oldest children. Six of the Fables assess the patients "fears and wishes". Of these six Fables 7, 9 and 16 are reliable for the youngest children, while Fables 13, 18 and 19 are reliable for the entire age range tested.

Sibling rivalry and parental rejection are presumably measured by two Fables each, Fables 6 and 20; and Fables 2 and 3 respectively. So far as reliability is concerned, Fables 6 and 3 meet the criterion in this study and, in fact, both are reliable for all the age groups. The other three variables or psychodynamics are associated with only one stimulus each. Concerning these the following observations can be made; Identification appears to be a reliable phenomenon with the intermediate age group only; castration fears are reliably assessed at all age levels but less so for the older than younger children; oedipal conflicts are expressed reliably by the youngest and oldest groups but more reliably by the oldest than the youngest group.

The meaning of greater or less reliability in a response in regard to conflict is almost impossible of interpretation in a study of this kind. Individual followup and intensive individual work would be necessary to reach any meaningful conclusions on this point. One might surmise that those Fables which elicit the greatest reliability for a particular age are assessing a more conflictual area, and are doing so with greater success than the other Fables. It is interesting to note that each of Fine's variables obtains significantly reliable responses from at least one age group. It would seem that the paucity of stimuli for certain of the variables may well be a more accurate explanation for this lack of reliability than the explanation that the variable is actually reliable with only one age.

Summary and Conclusions

The Fine revision of the Despert Fables was administered in group form to 177 school children in grades 3 - 8. The age span was 8 - 14 years. Six weeks later the same children were given the same test in the same way. Reliability was determined for each Fable by means of the Phi Coefficient. Fourteen or 70% of the Fables were found to have reliability at the 5% level or better for the entire group of subjects. Younger children were more reliable or consistent in their responses than were the older children. The results are discussed according to the psychodynamic formulation of each Fable as proposed by Fine.

From this study it may be concluded that 14 of the Fables give reliable results and that some Fables are reliable at certain ages and not at others. Thus, there is a different pattern of reliability according to the maturity of the subjects. If one feels that this kind of reliability is an important test characteristic, one would be justified in shortening the test and having different forms for different ages. In fact, the need for considerable revision of the test is indicated.

Bibliography

1. Despert, J. Louise. Psychosomatic Study of Fifty Stuttering Children. *Amer. J. Orthopsychiat.*, 1946, 16, 100-173
2. Dus, L. La Methods des Fables en Psychoanalyse. *Arch. Psychol.*, Geneva, 28, 1940.
3. Fine, Reuben. Use of the Despert Fables (Revised Form) in Diagnostic Work with Children. *Rorschach Research Exchange* and *J. Proj. Tech., 1948*, 12, 106-118.
4. Horst, Paul. Pattern Analysis and Configural Scoring. *J. clin. Psychol.*, 1954, 10, 3-11.

THE INTERDEPENDENCE OF RATINGS OF CASE HISTORIES*

WILLIAM A. DETERLINE AND A. W. BENDIG

University of Pittsburgh

PROBLEM

In the rating method, the task for the judge is to evaluate independently each stimulus against the constant standard of the descriptive phrases anchoring the rating scale categories, while in paired comparisons and ranking methods each stimulus is compared to the other stimuli in the set. This suggests that the ratings given to one stimulus will not be correlated with or influenced by the ratings given to the other stimuli, but this does not appear to be true in the practical rating situation. Guilford notes [8, pp. 322-323] that successive "absolute" judgments are often correlated and Bendig [1] found that the reliability of raters (average intercorrelation among the judges) is a positive function of the heterogeneity of the set of stimuli rated. Hamlin [9] concluded that the reliability and validity of clinical judgments are affected by characteristics of the stimuli presented and the total number of stimuli in the set. In rating clinical material, then, the rating given by a judge to one of the stimuli may influence and be influenced by his ratings of the other stimuli. If clinical ratings are interdependent, rather than independent, we might expect these dependency effects to increase as a function of the number of ratings elicited from each judge.

Several studies of the reliability of global ratings of clinical case material [2, 3, 4] have used the common "randomized blocks" experimental design [7, pp. 94-102] in which each of the judges (blocks) rates all of the cases (treatments). Others [6, 10] have used an "incomplete blocks" design [7, pp. 259-267] in which each of the judges rates only some of the total group of cases, with the assignment of particular combinations of cases to judges insuring that each judge rates the same number of cases and each case is rated by the same number of judges. Calvin [5] has recently developed a series of experimental designs that permit testing the two hypotheses concerning the interdependence of clinical judgments expressed above. Calvin's "doubly balanced incomplete blocks" designs place two restrictions on the selection of combinations of cases to be presented to the single judges: (a) each possible pair of cases is presented to the same number of judges, and (b) each possible triplet combination of three cases is presented to the same number of judges. Designs are given [5, pp. 83-88] in which the total number of cases ranges from 8 to 16 and the number of cases rated by each judge varies from 4 to 8. The analysis of variance model for these designs permits the isolation and testing for significance not only of the usual mean square for differences between treatments (cases) but also a mean square for correlation effects (interdependence between successive ratings by the same judge.)

The plan of the present study was to have three groups of judges make global judgments of the adjustment level of the behavior of clinical patients reported in abstracted case histories. A total of ten case histories was used and following the "doubly balanced incomplete block design" patterns each judge in the first group was given only four cases to rate, the second group received five cases while each judge in the third group rated six case histories. The significance of the interdependence measure was tested within each of the three groups of judges and the magnitude of this correlation effect was compared between groups to discover whether the interdependency between the ratings is a positive function of the number of cases rated by a judge. This measure of interdependence will be referred to as the "dependency-factor".

PROCEDURE

Stimuli. The stimuli used were ten abstracted case histories of male adults whose level of adjustment had been described by psychiatrists as ranging from relatively

*This paper is based on a thesis submitted by the first author to the Graduate School of the University of Pittsburgh in partial fulfillment of the requirements for the degree of Master of Science and was directed by the second author.

normal to psychotic. The case histories were one to two pages in length and have been used in several previous studies[2, 3].

Designs. From the series of designs presented by Calvin[5, pp. 84-88] three designs where the total number of stimuli is ten, were selected for use in this study. Table 1 lists the characteristics of each of the three designs. Each design specifies which of the ten stimuli are to be presented to a given judge and a separate booklet containing the appropriate combination of four, five, or six case histories was assembled for each of the 96 raters.

TABLE 1. SPECIFICATIONS OF DOUBLY BALANCED INCOMPLETE BLOCK DESIGNS

Description	Design A	Design B	Design C
Number of case histories	10	10	10
Number of case histories per rater	4	5	6
Number of raters	30	36	30
Number of ratings per case history	12	18	18
Number of times each pair is together	4	8	10
Number of times each triplet is together	1	3	5

Scale. The judges were asked to rate the adjustment level of the individuals described in the case histories in terms of a seven-point scale, verbally anchored with the phrases "slightly maladjusted", "moderately maladjusted," and "extremely maladjusted" at categories one, four and seven respectively. Categories two, three, five and six were unanchored. This scale is identical with those used in previous studies[2, 3, 4]. The scale was presented on a booklet face sheet which contained the instructions requesting the judges to make their ratings in terms of the scale only, to record each rating as they finished reading the individual case histories, and not to go back and change any early rating after reading later case histories.

Subjects. The judges were 96 undergraduate students in introductory, social and educational psychology classes with the *Ss* randomly assigned to one of the three rater groups. The case history booklets were distributed at one class session and the rating judgments collected at a subsequent class period. Previous studies[2, 3] have shown that the reliability of undergraduate *Ss* rating these case histories is comparable to that of graduate students rating the same cases and that their judgments correlate highly with psychiatric ratings.

RESULTS

The ratings were analyzed according to Calvin's analysis of variance procedure and the results for each of the three designs appear in Table 2. The methods previously used[2, 3, 4] were applied to the mean squares from the analysis of variance to obtain estimates of rater reliability and these estimates also appear in Table 2. Rater reliability is the average intercorrelation among the *Ss* in rating the cases.

TABLE 2. ANALYSES OF VARIANCE OF CASE HISTORY RATINGS.

	Design A			Design B			Design C		
Source of variation	d.f.	M.S.	F	d.f.	M.S.	F	d.f.	M.S.	F
Raters (unadj.)	29	3.32		35	2.03		29	3.46	
Case histories (adj.)	9	27.99	14.28*	9	53.69	47.09*	9	52.49	54.03*
Dependency Factor (adj.)	36	.95	.49	36	.94	.82	36	1.96	2.02*
Error	45	1.96		99	1.14		105	.97	
Rater Reliability			.52			.72			.75

*Significant at the .01 level.

The mean square for case histories was the only one that was significant in all three analyses, indicating that in each group the raters reliably discriminated different adjustment levels among the ten case histories.

The adjusted dependency factor variance is not significant in either design A or B (block size four and five respectively), but reaches a figure that is significant at the .01 level in design C (block size six). Table 2 indicates a positive relationship between number of ratings and the size of the dependency factor.

Individual rater reliability (interjudge reliability) became larger with increasing numbers of cases rated. Since this reliability measure is based on data which are clear of inter-rating dependency effects, it appears that reliability increases as a function of the additional case histories rated. Rater reliability for designs B and C was comparable in magnitude to that found when all ten cases were rated by each S[3].

Discussion

Our results indicate that clinical judgments are not independent, but show an interactive effect. Any single rating judgment is influenced by other similar judgments elicited from the S and there is some evidence that the interdependence of these judgments is a positive growth function of the number of similar judgments made. When four or five case histories were presented to the Ss the mean square for this dependency factor was not statistically significant, but did become significant when each S rated six cases. Calvin[5, p. 74] suggests that the interstimulus correlation effects may decrease when more stimuli are rated by each judge, but our results suggest that the opposite is true. It is easy to see how these effects could summate. If only two cases are rated by each judge then the rating of case number one is influenced only by the rating given case number two, and vice versa. However, if three cases are rated, then case one is influenced by both cases two and three. Apparently these case interactive effects are operative even when only a few cases are given each S, but only become important or statistically significant when more than five cases are rated by each judge.

It is interesting to note that rater reliability when five or six cases were rated by each S was slightly smaller or about the same size as those reported when all ten cases were rated by each judge[2, 3], but was lower when only four cases were judged. This suggests that there is a positive negatively accelerated relation between number of ratings per judge and rater reliability. Rater reliability apparently increases markedly between four and five cases per judge, increases slightly between five and six cases, and may remain constant or increase slightly between six and ten cases per S. This suggested relationship should be tested in another sample since our obtained rater reliabilities are influenced by considerable sampling error. It may be that any increased reliability with greater numbers of cases per judge is offset by the growth of the interstimulus dependency factor or by what Hamlin[9] calls "rater demoralization and fatigue."

One caution must be noted in generalizing our results. We used introductory psychology students as our judges and our conclusions as to the growth of the dependency factor in ratings may not apply to experienced clinicians. Previous research has indicated that undergraduate Ss are as reliable and valid in making global judgments of adjustment from case history material as are graduate students in psychology when longer rating scales are used[4] and when they rate the same set of case histories used in the present study[2].

Summary

Three groups of undergraduate students (N = 96) rated clinical case histories for adjustment level with each S rating either four, five, or six cases. The ratings were analyzed to yield measures of rater reliability and interstimulus dependency effects. Rater reliability was low (r = .52) for the judges rating four cases, but higher (r = .72 and .75) for those rating five or six cases. The interdependency among the ratings

was insignificant for the four and five case groups, but significant at the .05 level for the six-case Ss. It was concluded that clinical judgments are not independent, but each judgment is influenced by the other ratings made by the judge, and the correlation between the successive judgments made by an S is a positive function of the number of judgments elicited.

References

1. Bendig, A. W. Reliability of short rating scales and the heterogeneity of the rated stimuli. *J. appl. Psychol.*, 1954, 38, 167-170.
2. Bendig, A. W. Rater experience and case history judgments of adjustment. *J. clin. Psychol.*, 1955, 11, 127-132.
3. Bendig, A. W. The reliability of adjustment ratings and the length of case histories. *J. consult. Psychol.*, in press.
4. Bendig, A. W. and Sprague, J. L. Rater experience and the reliability of case history ratings of adjustment. *J. consult Psychol.*, 1954, 18, 207-211.
5. Calvin, L. D. Doubly balanced incomplete block designs for experiments in which the treatment effects are correlated. *Biometrics*, 1954, 10, 61-88.
6. Chambers, G. S. An investigation of the validity of judgments based on "blind" Rorschachs. Paper read at Midwestern Psychol. Ass., Chicago, April, 1955.
7. Cochran, W. G. and Cox, G. M. *Experimental designs.* New York: Wiley, 1950.
8. Guilford, J. P. *Psychometric methods.* (2nd ed.) New York; McGraw-Hill, 1954.
9. Hamlin, R. M. The clinician as judge: implications of a series of studies. *J. consult. Psychol.*, 1954, 18, 233-238.
10. Newton, R. L. The clinician as judge: total Rorschachs and clinical case material. *J. consult. Psychol.*, 1954, 18, 248-250.

A BENDER-GESTALT DIAGNOSTIC VALIDITY STUDY[1,2]

JOHN A. BOWLAND AND HERDIS L. DEABLER

University of Florida *VA Hospital, Gulfport, Miss.*

Introduction

Since Bender[(1)] published her monograph in 1938, the Bender-Gestalt (B-G) has been investigated from two main points of view. One approach has been to formulate tentative statements concerning its use as a qualitatively interpreted projective technique. The other main approach has been to establish the validity of an objective scoring system which is quantifiable and thus subject to statistical analysis. The present study utilizes both qualitative and quantitative ratings of initial and repeat B-G performances of normals, neurotics, schizophrenics and organics. The interdependent purposes of the investigation are as follows:

1. To determine the functional relationships between and within diagnostic groups, and between and within B-G initial and repeat performances as measured by Pascal-Suttell[(4)] scores.

2. To determine the diagnostic validity of B-G performances, utilizing both qualitative and quantitative ratings.

Procedure

The experimental group consisted of 80 male subjects including 20 "normals" and 60 subjects from a VA neuropsychiatric hospital. Group I, 20 normals, was composed of random selections from a population of 42 hospital employees and students in an adult education class. This population of 42 had been rated "normal" on the basis of screening test data evaluated by the staff psychologist supervising admission

[1]The authors wish to express their appreciation to Dr. Rayford T. Saucer, V. A. H., Gulfport, and Melvin Zax, University of Tennessee, for their pertinent suggestions concerning statistical analysis.

[2]From the Veterans Administration Hospital, Gulfport, Miss.

testing. Groups II, III and IV consisted of 20 neurotics, 20 schizophrenics and 20 organics for whom diagnosis by the psychiatric staff was made and which was in agreement with over-all psychological testing reports.

The B-G records, 2 for each subject, were the first and last tests of a 2½ hour admission screening battery administered to groups of 8 to 12. Group instructions were given, repeated individually when necessary, with each S having a set of B-G cards.

The records for all subjects were secured by a staff psychologist who did not participate in the rating or scoring. Each pair of records was assigned a code number and individual records were identifiable only as the initial or repeat performance. The pairs were stapled together, randomized and presented to each judge in the same order.

The clinical or subjective ratings were made by seven staff psychologists and advanced psychology trainees. Each rater was provided with a list of 80 coded numbers, corresponding to the order of presentation of pairs of records, wherein he could check "normal," "neurotic," "schizophrenic" or "organic." Instructions were to consider each pair as independently as possible, *i.e.*, the judge was neither to rank them nor distribute them into four piles, nor to consider total number in each group.

Each record was also scored quantitatively by the Pascal-Suttell method under the same condition of anonymity. The raw scores for each pair of records were averaged and the resultant 80 mean scores ranked. These ranked scores were in turn divided into quartiles. The first quartile (the 20 "best" scores) were rated as "normal," the second as "neurotic," the third as "schizophrenic" and the fourth as "organic."

Purpose 1, determining the relationships between and within the various groups and between and within the initial and repeat B-G performances, was carried out by a complex analysis of variance[3].

Purpose 2, determining the diagnostic predictability for each method of rating, was carried out by calculating a *T* ratio between the percentage of mean correct ratings obtained by the seven judges and the expected percentage, and between the percentage of correct ratings obtained by the quantitative method and the expected percentage.

Results

1. The diagnostic groups' performances, in raw Pascal-Suttell scores, reflected a rank-order correlation such that the more severe the disorder, the higher the mean and the greater the variance. Table 1 shows the means and variances for each group, computed from the average of initial and repeat test deviation scores. Since this correlation was likely to violate the condition of independence of mean squares, necessary for an analysis of variance, the data were subjected to a square root transformation.[3] The results of the complex analysis of variance are shown in Table 2. The two rows represent Test 1 and Test 2 scores. The four columns represent the scores of Groups I-IV.

Table 1. Deviation Scores on Test 1 and Test 2 by Groups I-IV in Means and Variances

Group	Number	Means	Variance
I Normals	20	16.4	174.0
II Neurotics	20	28.1	315.4
III Schizophrenics	20	42.8	497.0
IV Organics	20	80.0	3446.6

[3]While the square root transformation of data is statistically valid, there is some question as to how meaningful manipulations of transformed data are for routine clinical use. A non-parametric analysis of the raw data was incidentally performed[2]. The results, available on request, are entirely consistent with those of the analysis of variance.

TABLE 2. ANALYSIS OF VARIANCE

Source of Variance	Sum of Squares	DF	Mean Square	F
Rows	.26	1	.26	—
Columns	522.96	3	174.32	60.1[4]
Interaction	.39	3	.13	—
Within	438.73	152	2.88	
Total	961.34	159		

[4]Significant beyond .001.

Table 2 demonstrates the columns to be the major source of variance. Since the interaction variance was smaller than the within variance, the latter was used as the error term. The ratio of column variance to within variance, with 3 and 152 degrees of freedom, yields an F of 60.1, significant beyond the probability level of .001. Neither rows nor interaction are significant sources of variance. These results mean that the Null hypothesis of no true differences between the means of Group I-IV may be rejected at a high level of confidence. The group performances on Test 1 did not differ significantly from those on Test 2.

2. Of a total of 80 ratings made by the Pascal-Suttell scoring system, means of initial and repeat tests ranked in quartiles, the obtained number of correct diagnoses was 40, compared with an expected chance frequency of 20; *T* equals 3.5. The 7 judges' mean number of correct ratings was 39.6, compared with the same chance frequency of 20 out of 80; *T* equals 3.5. The frequency of correct ratings by either method exceeds the chance expectation beyond the probability of .001.

The mean performance of normals, neurotics, and schizophrenics are comparable to the performance of Pascal and Suttell's normative groups of (high school) non-patients, psychoneurotics and psychotics. The mean raw scores differ only 1-5 points. The differences in performance obtaining from individual and group administration of the B-G appear to be negligible.

SUMMARY

1. One hundred and sixty B-G tests, two for each of 20 normals, 20 neurotics, 20 schizophrenics and 20 organics, were obtained. They were the first and last tests of a two and one-half hour group screening battery.

2. An analysis of variance was computed to determine the differences between and within diagnostic groups, between and within the initial and repeat performances. The differences between groups were significant at a high level of confidence. The quantitative differences between initial and repeat tests were not statistically significant.

3. Each pair of B-G tests was evaluated as the performance of a normal, schizophrenic or organic by qualitative and quantitative methods. Either method obtained successful diagnoses significantly better than chance.

REFERENCES

1. BENDER, LAURETTA. *A visual motor Gestalt test and its clinical use.* New York: Amer. Orthopsychiat. Ass., 1938.
2. KRUSKAL, WILLIAM H., and WALLIS, W. ALLEN. Use of ranks in one criterion variance analysis. *J. Amer. Statist. Ass.*, 1952, *47*, 583-621.
3. MCNEMAR, QUINN. *Psychological statistics.* New York: John Wiley & Sons, 1949.
4. PASCAL, G. R. and SUTTELL, Barbara J. *The Bender-Gestalt test.* New York: Grune and Stratton, 1951.

RORSCHACH CARD REJECTION IN NORMAL AND PSYCHIATRIC GROUPS[1]

BOYD D. SISSON
University of Nebraska

EARL S. TAULBEE
VA Mental Hygiene Clinic, Omaha and University of Nebraska College of Medicine

AND

CHARLES O. GASTON
University of Houston

Problem

Even with the voluminous research which has been devoted to the Rorschach test, there has been little more than speculative or indirect reference to the significance of Rorschach card rejection by subjects. In the main, these considerations of rejection have been products or observations of grosser Rorschach test appraisals, yet rejection is attributed interpretative significance. For example, Beck[(2)] states that individuals will show trends to rejection because of their fearful needs for certainty; however, they seldom reject as the fear of not responding is just as cogently threatening. Klopfer and Kelley[(4)] emphasize the significance of rejection in terms of the specific card(s) rejected. On the other hand, Brussel et al.[(3)] state: "Rejection of cards: This does not occur in normal adults. Rejection is highest in schizophrenic cases and may happen in any card. Rejection of colored cards—and on occasion of highly-shaded cards—occurs in the psychoneurotics. Mental defectives may reject the more complicated cards such as IX. Any rejection is cause for suspicion." Miale and Harrower-Erickson[(6)] report from their study of 43 neurotics and 20 normals that refusals to one or more cards occurred in 47% of the neurotics but was absent in the normals.

One study dealing specifically with Rorschach card rejection is a recent one conducted by Mensh and Matarazzo[(5)], who note the scarcity of systematic investigations related to this topic and report data on rejection in psychodiagnosis and its relationship to other personality characteristics. The present study was designed partly as a cross-validation of Mensh and Matarazzo's findings, and, in addition, to extend the data on this aspect of Rorschach test behavior by including a sizeable group of normal adults as well as normal and mentally defective children.

Study

The subjects consisted of 190 normal, 72 neurotics, and 128 schizophrenic adults, and 200 normal and 44 mentally defective children. The neurotics and schizophrenics were VA hospital inpatients diagnosed by the psychiatric staff; the mental defectives were institutionalized; the normal subjects were residents of a small midwestern community[2] located in the same geographical locale from which the patients came.

The Rorschach test was individually administered to all subjects, following the procedure of the Beck system[(1)], and the protocols were analyzed to determine the occurrence of card rejections.

Results

The normal and mentally defective groups consisted of both males and females, and analysis to determine sex differences with respect to the number of individuals rejecting one or more cards revealed no significant difference. Therefore, the data for males and females were combined. Table 1 presents the subject groups, identifying information, and rejection information.

---

[1]From the Veterans Administration Hospital, Omaha, Nebraska.

[2]The records were from a study conducted under the supervision of Dr. Marshall R. Jones and supported by the University of Nebraska Research Council.

TABLE 1. SUBJECTS, AGE, RESPONSE TOTAL, AND NUMBER AND PERCENTAGE OF CASES REJECTING RORSCHACH CARDS

Groups	N	Age		R		Cases Rejecting	
		Mean	SD	Mean	SD	No.	%
ADULTS							
Normals	190	39.20	10.89	27.40	16.24	33	17.4
Males	83						
Females	107						
Neurotics	72	32.80	6.26	22.30	10.71	24	33.3
Schizophrenics	128	29.86	6.21	25.66	18.75	41	32.0
CHILDREN							
Normals	200	12.63	3.12	26.82	12.45	20	10.0
Males	90						
Females	110						
Mental Defectives	44	12.90	2.26	13.60	4.58	10	22.7
Males	22						
Females	22						

Chi square tests of differences[3] between the number of cases rejecting in the different groups reveal significant differences between the normals and neurotics ($X^2 = 7.81$, $P < .01$) and the normals and schizophrenics ($X^2 = 9.21$, $P < .01$). However, the difference between neurotics and schizophrenics did not approach a level of conventional significance. Within the children's group, the number of cases rejecting among the mental defectives is significantly greater than the number among the normal children ($X^2 = 5.42$, $P < .02$). Arbitrary division of the normal children into those 12 years of age and below as compared with those 13 years and above does not show a significant difference for number of cases rejecting.

As can be seen in Table 1, the groups are not matched for age and number of Rorschach responses. In an effort to control possible variance due to these two factors, subgroups for each of the adult classifications—117 normals, 65 neurotics, and 100 schizophrenics—were selected from the total groups for the purpose of obtaining matched samples on mean age and response total, eliminating those records with less than 10 responses. Further, based on education data of 43 of the neurotics and 82 of the schizophrenics they were matched for educational level. Wechsler-Bellevue Intelligence Scale scores on 20 of the neurotics and 67 schizophrenics show them to be matched on IQ. The mean education (neurotics 10.67; schizophrenics 11.67) and mean IQs (neurotics 105.35; schizophrenics 103.52) suggest as well that the normals are probably similar in these respects based on average normal education and IQ expectancy.

Analysis of the individual card rejections for the unmatched groups reveals that schizophrenics reject cards V, VII, VIII, IX and X, and neurotics card IX, significantly more often than normals (Table 2). However, these differences, with the exception of card X between the schizophrenics and normals, were not found among the matched groups. The unmatched groups were compared with respect to the number of individuals rejecting zero or one card versus two or more and zero to two cards versus three or more. The findings indicate that in both instances the schizophrenics rejected significantly more cards than did the normals; the neurotics rejected more than the normals only when considering the number of cases rejecting zero or one card versus two or more. The schizophrenics did not reject significantly more than the neurotics in either of the comparisons. When the same analysis of the data is made using the matched groups there are no significant differences. Thus it can be seen that when the groups are matched the only significant difference re-

[3]In all instances where a cell frequency was five or less Fisher's "exact method" was employed.

TABLE 2. NUMBER AND PERCENTAGE OF INDIVIDUAL CARD REJECTIONS

Groups	I		II		III		IV		V		VI		VII		VIII		IX		X	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
ADULTS																				
Normals	0	0.0	3	1.6	4	2.1	5	2.6	7	3.7	9	4.7	7	3.7	1	.5	12	6.3	8	4.2
Neurotics	0	0.0	2	2.8	2	2.8	5	6.9	5	6.9	7	9.7	6	8.3	0	0.0	11	15.3	8	11.1
Schizophrenics	1	.8	7	5.5	4	3.1	7	5.5	13	10.2	10	7.8	16	12.5	5	3.9	19	14.8	19	14.8
CHILDREN																				
Normals	0	0.0	2	1.0	2	1.0	9	4.5	3	1.5	3	1.5	6	3.0	2	1.0	7	3.5	2	1.0
Mental Defectives	3	6.8	4	9.1	2	4.5	5	11.4	1	2.3	3	6.8	4	9.1	2	4.5	5	11.4	3	6.8

maining with respect to all rejection data is the more frequent rejection of card X by schizophrenics than by normals.

The Rorschach tests were administered to the majority of the normal subjects by three examiners, two males and one female, and to the majority of the pathological groups by three examiners, including two of the authors (B. S. and E. T.). No significant differences were found among the six examiners with respect to the number of their cases rejecting.

Discussion

The above findings, essentially in agreement with those reported by Mensh and Matarazzo, are viewed by the authors as evidence that Rorschach card rejection has little or no diagnostic significance in terms of the common nosological disorder categories. We feel that the inclusion in our samples of sizeable normal adult and children's groups particularly emphasizes this view. These findings, however, do not suggest that rejection is without behavioral significance. It would appear that card rejection is more likely to be related to personality variables which cut across the broad diagnostic classifications.

Summary

The Rorschach protocols of 634 subjects, representing groups of normal, neurotic and schizophrenic adults, and normal and mentally defective children, were analyzed for the occurrence of card rejections. For the adult population no significant differences in number of cases rejecting were obtained in the matched groups. However, in the unmatched samples the normals rejected significantly less often than did the neurotics and schizophrenics, although the latter two groups did not differ significantly from each other. Among the children, significantly more mental defectives rejected cards than did their normal age-peers.

The only significant difference for individual card rejection obtained in both the matched and unmatched groups occurred between the normal and schizophrenic adults on card X, with more instances of rejection among the latter group. Comparison of six examiners showed no significant examiner variation with respect to number of their cases rejecting.

The findings are viewed as evidence that the phenomenon of Rorschach card rejection is of little value in terms of broad psychodiagnostic differentiation. The results do not preclude the possibility that there is personality significance in card rejection behavior.

References

1. Beck, S. J. *Rorschach's test. Vol. I—Basic Processes.* New York: Grune & Stratton, 1949.
2. Beck, S. J. *Rorschach's test. Vol. III—Advances in interpretation.* New York: Grune & Stratton, 1952.
3. Brussel, J. A., Hitch, K. S., and Piotrowski, Z. A. *A Rorschach Training Manual.* Utica, New York: State Hospital Press, 1950.
4. Klopfer, B., and Kelley, D. M. *The Rorschach technique.* Yonkers-on-Hudson, New York: World Book Company, 1942.
5. Mensh, I. N., and Matarazzo, J. D. Rorschach card rejection in psychodiagnosis. *J. consult. Psychol.*, 1954, 18, 271-275.
6. Miale, F. R. and Harrower-Erickson, M. R. Personality structure in the psychoneuroses. *Rorschach Res. Exch.*, 4, 1940, 71-74.

A COMPARISON OF THE ABILITY OF NORMAL AND BRAIN INJURED SUBJECTS TO PRODUCE ADDITIONAL RESPONSES ON A SECOND ADMINISTRATION OF THE RORSCHACH TEST*

LEONARD WEINER AND EARL BROWN
University of Kansas *Wadsworth V. A. Hospital*

BERT KAPLAN
University of Kansas

Problem

In a recent study, Kaplan and Berger[2] demonstrated that a group of college students were able to produce large numbers of Rorschach responses in addition to those given on a first administration. These workers gave the test, in a group situation, four times, but on the second, third and fourth administrations their subjects were asked to report only new responses and to omit mention of any they had given before. The first administration produced an average of 34 responses, the second 24, the third 21 and the fourth 24. The conclusion was that the initial Rorschach responses of normal subjects constitute only a sample of a much larger universe of responses which they are capable of giving. In interpreting their results, Kaplan and Berger raise two questions, first whether their findings can be generalized to populations other than college students who are noted for their high productivity on the Rorschach, and second whether the number of responses produced on a first administration of the test is really related to capacity.

The present study sought to discover whether individuals with an incapacity which is known to be related to the ability to give many Rorschach responses would be able to produce substantial numbers of new responses on a second administration. The subjects chosen had organic brain injury and it was predicted that in view of this group's difficulty in shifting and reorganizing perceptual stimuli, they would not be able to do so. This result would help to anchor the phenomenon described by Kaplan and Berger by demonstrating that the ability to give additional Rorschach responses is not independent of capacity. In addition, there was the possibility that the new procedure would be of help with the difficult diagnostic task of distinguishing normal individuals from those with organic brain impairment. Finally, the performance of the control group of normal subjects would enable us to tell whether the phenomenon under study is limited to populations of high intellectual levels or exists at other levels as well.

Procedure

Two groups of ten subjects each were used. The experimental group consisted of individuals who had suffered an organic injury to the brain at least six but not more than 24 months prior to testing. Each prospective subject was administered the Trail Making Test[1], a test which in the words of Partington and Leiter[4] involves "integration of previously learned material into a relatively complex pattern of response requiring shifts and fluctuations in attention, . . . retention, recall, organization and synthesis". Only those subjects who were unable to complete the test without error were included in the experimental group. Two of the subjects selected for the experiment passed the test and were not used. The mean age of the subjects in the group was 45.6 years and the mean number of years of education was 12.2. The control group was composed of normal individuals and included eight fracture cases, 1 poliomyelitis case and 1 arthritis case. This group was matched to the experimental group in age and education. All of these subjects were able to pass the Trail Making Test without error. All subjects were male and participated on a voluntary basis. Of those solicited only two individuals refused to participate.

*The data for this study were collected with the cooperation of members of the staff of Wadsworth V. A. Hospital, Leavenworth, Kansas.

Each subject was presented the ten cards in the usual manner with the instructions following those recommended by Klopfer[3]. However, each subject was asked to retain the card for at least 60 seconds and after the first three cards was asked if he could see anything else. The inquiry was altered in that the subjects were asked to indicate the locations of their responses only. They were told to feel free to add new responses and asked specifically after each card if they could see anything new. No additional responses were forthcoming in either group. Two days later the subjects were recalled and told that they would be given the test a second time, and that this time they were to attempt to give as many new and different responses as possible. This instruction was repeated prior to the presentation of each card.

Results

The results of the experiment in Table 1 indicate that normal and experimental subjects produced on the average almost the same number of responses on the initial administration. On the second administration, however, the control group had an increment of 86%, whereas the experimental group was able to produce an increment of only 17%. The mean number of responses for the control group was 11.6 on the first administration and 10.0 on the second. For the experimental group, the mean number of responses was 11.7 on the first administration and 2.0 on the second. The difference between the means of responses on the second administration in the two groups is significant at the .001 level.

Table 1. The Number of Responses by the Matched Individuals on the First and Second Administrations of the Rorschach Test

Pairs of S's	Group C Admin. I	Group C Admin. II	Group E Admin. I	Group E Admin. II
1	10	10	10	1
2	9	8	6	1
3	11	13	9	2
4	10	13	17	2
5	21	14	6	2
6	11	11	20	6
7	18	13	20	3
8	9	7	11	1
9	9	2	8	3
10	8	9	10	0
Total Responses	116	100	117	20
X's:	11.6	10.0	11.7	2.0

Summary and Conclusions

The results indicate that brain injured patients were able to produce only a few additional responses on a second administration of the Rorschach in contrast to the normal individuals who produced many new ones. This suggests that capacity can impose sharp limits on Rorschach responsivity, at least when the incapacity is well defined and unusually low. While it is not inconceivable that with altered experimental conditions and increased motivation, the brain injured subjects could produce some additional responses, the relative lack of variation in the number produced on the second performance suggests that capacity has been a limiting factor. It may be questioned whether this is a very surprising conclusion. However, in the light of Kaplan and Berger's findings that there is, in normal college subjects, a very great disparity between performance and capacity, it is important to know that in the brain injured patients, at least, there is not this disparity.

In addition, the substantial productivity of the control group on the second performance indicates that this phenomenon is not limited to groups with specially

high ability or motivation. It suggests instead that even where clinicians expect to find very low Rorschach productivity, as in hospital populations, the capacity exceeds performance and hope is therefore offered that by a more appropriate structuring of the test situation, richer and more numerous responses might be elicited and the test could become more useful.

Finally, the procedure seems to offer some promise as a diagnostic tool. An examination of the Rorschach records themselves reveals that while a few of the test records of the brain injured patients are readily recognizable as such, it is difficult or impossible to distinguish the majority of cases from the normal records. The second performances, however, can be very clearly distinguished simply in terms of the number of new responses produced. Of the 20 subjects in the study, only two could not be placed in the correct group by this criterion. It appears to be much easier to correctly diagnose brain injury on the basis of the second performance on the Rorschach Test than on the first performance. The present study is, of course, only suggestive. Many more subjects of more heterogeneous types are needed before it is known whether this procedure will prove useful in actual clinical work.

References

1. Armitage, S. G. An analysis of certain psychological tests used for the evaluation of brain injury. *Psychological Monographs*, 1946, vol. 60.
2. Kaplan, B. and Berger, S. Increments and consistancy of performance in four repeated Rorschach administrations. Paper read at the 1955 meetings of the Midwestern Psychological Association.
3. Klopfer, B. and Kelley, D. *The Rorschach Technique*. Yonkers-on-Hudson: World Book Co., 1946.
4. Partington, J. E. and Leiter, R. G. Partington's Pathway test. *Psychological Service Center Bulletin*, 1949, 1.

GALVANIC SKIN CONDUCTANCE AS A FUNCTION OF SUCCESSIVE INTERVIEWS

BARCLAY MARTIN

University of Wisconsin

As part of an exploratory study designed to assay the possibilities of a more controlled approach to the investigation of psychotherapy, galvanic skin conductance measures were obtained in successive interviews with the same subjects. This report is concerned with the findings in regard to these skin conductance measures. Previous findings based on the analysis of individual records by Light [7] and Bixenstine [2] suggest that palmar sweating as measured by a colorometric procedure may be a reasonably valid index of tension changes during the process of psychotherapy. Other findings outside the context of psychotherapy also suggest that changes in skin conductance as measured by the more conventional GSR apparatus may be useful as an indicator of changes in tension level, activity level or "anxiety," the particular label employed varying with the experimenter [3, 4, 5, 8]. As was the case with the above mentioned studies the interest in this research was in the base level of skin conductance rather than the transient changes known as the galvanic skin reflex.

[1]Technically the term conductance is not appropriate here since the reciprocal of the resistance taken out of the bridge circuit to balance the S does not directly reflect the pure resistance of the skin but is also affected by reactance due to capacitance since AC current is used. The measurement as obtained is operationally defined though, and this inaccuracy in terminology from the point of view of electrical engineering probably does not affect the pyschological meaning of the results.

Procedure

Two undergraduate summer school classes of 86 and 52 students respectively were given the Taylor Manifest Anxiety Scale during their regular class hour. In addition E asked that any student who would be interested in participating in five successive "interview-like" sessions to so indicate by placing a particular mark on the test answer sheet. All Ss who scored 22 or higher on the Taylor Anxiety Scale and who volunteered for the "interview-like" sessions were seen for an initial session by E. In this initial session finger electrodes were put on the S and readings of galvanic skin conductance taken at specified times. The apparatus used to measure skin conductance was an AC circuit designed by Grant[6]. Dry stainless steel finger electrodes were used. The session began with five minutes of relaxation followed by the Wechsler-Bellevue memory for digits forward item, followed by a 15-minute interview, followed by 15 minutes of free association, followed by five minutes of relaxation. The interview proceeded in a relatively nondirective fashion after an initial request for S to tell something about himself. In the free association period S was asked simply to say what came to mind no matter how trivial, embarrassing, etc., the thought might be; in other words, the classical free association instructions. At the end of the session the S was asked if he would like to come back for four more sessions similar to this one, now that he had a better idea of what was involved. All indicated that they would like to do so.

Pairs of Ss were then matched as closely as possible on the following measures: (1) Taylor Anxiety Scale, (2) log skin conductance in first five minutes of relaxation, and (3) the maximum change in log skin conductance during the memory for digits task from the base level of the initial relaxation period. This latter measure was intended to give some indication of general GSR reactivity. It was possible to obtain 10 fairly closely matched pairs which were then split randomly into experimental and control groups. The means and sigmas for the two groups are shown in Table 1. The experimental group was composed of 8 women and 2 men, the control group of 7 women and 3 men.

Table 1. Means and Sigmas for Experimental and Control Groups

Measures	Experimental		Control	
	Mean	Sigma	Mean	Sigma
Taylor Anxiety	29.2	4.0	29.5	4.1
Log conductance, relaxation	1.333	1.116	1.323	1.081
Change log conductance, memory for digits	1.123	1.043	1.136	1.090

Note: Conductance is measured in micro-mhos.

Each S was now seen for 4 more sessions. The sessions for the 10 experimental Ss were conducted as follows: 5 minutes relaxation, 15 min. free association, 15 min. discussion, 5 min. relaxation. For the 10 control group Ss the sessions were as follows: 5 min. relaxation, 15 min. free association, 15 minutes listening to own associations played back to them, 5 minutes relaxation. Readings of galvanic skin conductance were taken every minute during relaxation periods and every five minutes during free association period, discussion period and listening to own association period.

Results

The results in regard to galvanic skin conductance are illustrated in Fig. 1. The unit of measurement used was log conductance so that decreasing curves indicate increasing resistance or psychologically decreasing "tension."

Alexander's method of trend analysis[1] was applied to the curves obtained for the free association periods, and group means were found to be significantly different,

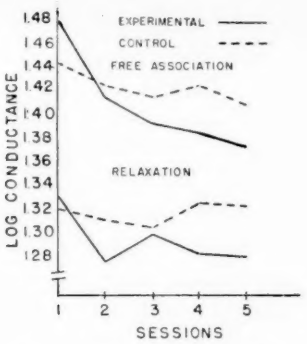

FIG. 1. SKIN CONDUCTANCE AS A FUNCTION OF SUCCESSIVE SESSIONS.

$p < .01$, and the overall slope of both groups combined was significant, $p < .05$; however, the difference between group slopes was not significant. The difference between group means should be interpreted cautiously since it is apparent that the experimental group decreased considerably more than the control group between the initial session and the first session with no difference in treatment. The tendency for skin conductance to decrease as a function of successive sessions is, however, a relatively unambiguous finding, and is even more striking if the initial session is included in the curves.

The curves for skin conductance during the beginning relaxation period do not show the same tendency to decrease as a function of successive sessions as those for the free association period. Neither the overall group slope is significantly different from zero, nor are the individual group slopes significantly different from each other. The means of the two groups are significantly different but this cannot be attributed to differential treatment since the groups are quite different at the beginning of the first experimental session. In other words, although they were matched fairly closely on the basis of the initial session they changed to such an extent before the first differential treatment session as to render them no longer matched on this measure.

An estimate of the stability of the skin conductance measure over the four free association periods was obtained by an analysis of variance procedure in which the measures were considered as four replications for each of 20 Ss. The reliability of a score based on the average skin conductance for all four sessions was computed to be .95. This indicates that the Ss were highly consistent in their relative position in the distribution of skin conductance scores from session to session even though the group means dropped significantly as a function of successive sessions.

SUMMARY

Twenty Ss who scored high on the Taylor Anxiety Scale were seen for five successive sessions. Ten of the Ss, the experimental group, experienced a free association period followed by a discussion of their associations; and the other ten, the control group, experienced a free association period followed by listening to their own associations played back to them. Measures of galvanic skin conductance were obtained during all sessions. For the free association periods there was a significant tendency for the average skin conductance to decrease as a function of successive sessions for both groups. However, there was no difference between the trends for the experimental and control groups. Average skin conductance during an initial relaxation period did not change as a function of successive sessions, but was less than the skin conductance during free association for all sessions.

These findings are taken as indicating the possible usefulness of skin conductance as a measure of change in general tension level, and more specifically in this instance as indicating the tendency of a tension or apprehension response to extinguish as a function of successive sessions.

References

1. Alexander, H. W. A general test for trend. *Psychol. Bull.*, 1946, *43*, 533-557.
2. Bixenstine, V. E. A case study of the use of palmar sweating as a measure of psychological tension. *J. abnorm. soc. Psychol.*, 1955, *50*, 138-143.
3. Coppock, H. W. Responses of subjects to their own galvanic skin responses. *J. abnorm. soc. Psychol.*, 1955, *50*, 25-28.
4. Davis, R. C. Modification of the galvanic reflex by daily repetition of a stimulus. *J. exp. Psychol.*, 1934, *17*, 504-535.
5. Duffy, E. and Lacey, O. L. Adaptation in energy mobilization: changes in general level of palmar skin conductance. *J. exp. Psychol.*, 1946, *36*, 437-452.
6. Grant, D. A. A convenient alternating current circuit for measuring GSR's. *Amer. J. Psychol.*, 1946, *59*, 149-151.
7. Light, B. H. Tension changes in patients undergoing psychotherapy. Unpublished Ph.D. thesis, University of Illinois, 1951.
8. Schlosberg, H. Three dimensions of emotion. *Psychol. Rev.*, 1954, *61*, 81-88.

MANIFEST ANXIETY IN PSYCHOTICS

HAROLD RUBIN LEO SCHNEIDERMAN
AND
WILLIAM C. HALLOW ROBERT J. JONES

Veterans Administration Hospital, Lebanon, Pa.

Problem

The present study is an attempt to evaluate manifest anxiety in psychotics in terms of its relationship to level of hospital adjustment. Previous studies by Mowrer [3] and Meadow, Greenblatt, Levine and Solomon [2] have indicated that some relationship exists between levels of adjustment and expressions of comfort or discomfort. In the study by Meadow et al. [2] with psychotic subjects it was found that a relatively high degree of expressed discomfort (as measured by DRQ) was a valid indicator of good adjustment. The purpose here is to further explore this suggested relationship by comparing results of the Taylor Manifest Anxiety Scale [5] with level of hospital adjustment. Within the framework of the present investigation the MAS is regarded as providing a measure of symptoms of physical and psychological discomfort as reported by the subject. The MAS, unlike the DRQ, may be viewed as a measure, in large part, of symptoms of a more behavioral nature, (restlessness, irritability, etc.), capable of interfering with adjustment. Presumably, MAS scores also reflect both the subject's ability and willingness to express such symptoms or complaints, as has been demonstrated by Buss, et al. [1] As a measure of adjustment level, the Hospital Adjustment Scale (HAS) [4] was selected in terms of its suitability for evaluating psychotic patients. By means of this scale, three different aspects of behavior can be evaluated: I. Communication and Interpersonal Relations, II. Care of Self, and III. Social Responsibility, and Work Activities and Recreation.

Subjects and Procedures

Forty-five chronic, hospitalized, functional psychotic patients, largely WW II veterans, were used as subjects. These subjects are part of a research project on reserpine to be reported at a future date. The MAS was administered individually

to each subject as part of a psychological test battery. The HAS ratings were made for all patients by a psychiatric aide of the ward who had observed them for at least a three-month period.

Results and Discussion

Test-retest values were computed for 15 randomly selected subjects. The reliability coefficient for HAS was found to be .96, and that for MAS .91. Comparison of test-retest means for HAS and MAS respectively, failed to reveal any significant change when a *t*-test was applied. These data indicate the high degree of stability for these measures over a one-month period.

Table 1. Intercorrelations Between MAS and HAS (N = 45)

Measures	HAS I	HAS II	HAS III	HAS Total
MAS	−.50*	−.46*	−.38*	−.49*
HAS I		.77*	.78*	.95*
HAS II			.66*	.87*
HAS III				.84*

*Significant at .01% level or better.

Table 1 contains the obtained intercorrelations between the MAS and HAS scales. As can be observed, a significant negative relationship ($r = -.49$) was found between MAS and total HAS. This would indicate that the patients who were rated high in adjustment on the HAS scale tended to express relatively few symptoms of manifest anxiety. Further analysis of the data (see Table 1) revealed that this relationship obtained, as well, for the three sub-scales of the HAS. Inspection of the frequency distribution for total HAS showed a tendency toward bimodality, with some tendency toward skewness at the upper end of the scale. This finding raised a question concerning the role of *level* of hospital adjustment in evaluating manifest anxiety scores. In other words, would a breakdown of the HAS distribution into high and low scoring groups provide additional information concerning the relationship between MAS and HAS? Accordingly, a Chi-square analysis of HAS and MAS scores was made, dividing the distributions of scores of the variables into high and low on the basis of their respective mean scores. Table 2 contains the X^2 values thus obtained.

Table 2. X^2 Values Between HAS Variables and MAS (N = 45)

Measures	MAS (Mn = 12) X^2	P
HAS I Mn = 64.95	3.09	<.10
HAS II Mn = 69.49	9.01	<.01
HAS III Mn = 68.07	8.60	<.01
HAS Total Mn = 67.22	5.66	<.02

The means for total HAS and MAS were respectively 67.22 and 12.07. A Chi-square value of 5.66, significant at the .02 level of confidence, was obtained. The distribution of scores on a fourfold table revealed the greatest frequency (N = 20) in the cell designating high adjustment /low anxiety. The low adjustment group were

fairly evenly distributed between high and low scoring groups on the MAS. These findings provide an answer to the question raised above concerning the relationship between level of hospital adjustment and manifest anxiety scores. It appears that patients with higher levels of adjustment tend to express relatively fewer symptoms of manifest anxiety, while the more poorly adjusted patients seem to be much less predictable in this respect. The members of the latter group, apparently, may or may not express many signs of discomfort, probably depending upon their current mental condition. That is to say, the poorly adjusted patient may be apathetic, or free of subjective discomfort; or he may experience some degree of anxiety, but be unable or unwilling to express it; or, finally, he may both experience and express subjective discomfort, ranging from very mild to very extreme degrees of anxiety. As to why the well adjusted group expresses relatively little anxiety, two possible explanations present themselves. The first, and most obvious answer is that its members actually experience less discomfort or anxiety than those of the poorly adjusted group. The second possibility is that they have learned to adjust themselves to our hospital program, and its emphasis upon conformity, by suppressing their complaints and passively going along with the prescribed routine.

Chi-square analyses of the three sub-scales of the HAS (Table 2) revealed the same kind of relationship to MAS in each case as that obtained for the total HAS scale. Well adjusted patients expressed fewest signs of distress. It should be added that a slight qualification holds for HAS I which deals with "Interpersonal Relationships." The obtained X^2 value of 3.09 is significant between the 5% - 10% levels of confidence. At the least, this indicates a definite trend in the expected direction.

To determine which of the three HAS sub-scales contributed most to the total HAS, intercorrelations were computed for the various categories. Table 1 contains this information. These correlations very closely approximate those reported by the authors of the test [4]. It was found that the subscale dealing with "Communication and Interpersonal Relationships" (which also contains the largest number of items in the scale) contributes most heavily to the final adjustment rating. It would seem, then, that ratings on Scale I were more representative of the overall adjustment rating than either of the other sub-scale ratings.

Summary

The present study attempted to evaluate the role of level of hospital adjustment in relation to manifest anxiety. Forty-five hospitalized psychotics were tested with the Hospital Adjustment Scale and the Taylor Manifest Anxiety Scale. The findings appear to indicate strongly that level of hospital adjustment should be taken into consideration in evaluating the manifest anxiety of psychiatric patients, at least of psychotics. Psychotics who are well adjusted to the hospital program seem to be relatively free of expressed manifest anxiety symptoms. Those who are poorly adjusted may express much or little manifest anxiety. Possibly failure to control for level of adjustment has resulted in some of the inconsistencies reported in studies using the MAS as a criterion measure.

References

1. Buss, A. H., Wiener, M., Durkee, A., and Baer, M. The measurement of anxiety in clinical situations. *J. consult. Psychol.*, 1955, 19, 125-129.
2. Meadow, A., Greenblatt, M., Levine, J., and Solomon, H. C. The discomfort-relief quotient as a measure of tension and adjustment. *J. abnorm. soc. Psychol.*, 1952, 47, 658-661.
3. Mowrer, O. H. *Psychotherapy theory and research.* New York: Ronald, 1953.
4. McReynolds, P. and Ferguson, J. T. *Clinical manual for the Hospital Adjustment Scale.* Stanford U. Press, 1953.
5. Taylor, Janet A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol., 1953, 48,* 285-290.

EDITORIAL OPINION

PROFESSIONAL IMPERIALISM

Professional imperialism may be defined as the efforts of any one professional group to gain working control over any field to the exclusion of other potentially competing groups, *i.e.*., to carve out an empire for themselves to the possible detriment of the whole field. This is not a new phenomenon since nature seems to abhore a vacuum, and wherever power or material resources exist in a free state there inevitably develops a struggle among humans to control them. Professional imperialism reflects the operation of the same political forces which have carved out empires in the evolution of human civilization. In some instances, imperialistic forces have been benevolent and constructive but too often they have been characterized by a selfish exploitation of the masses for the benefit of the few. Usually there is an attempt to rationalize such imperialism by claiming that it is accomplished only in the best interests of all concerned. Usually, however, ulterior motives may be discerned in the attempt to garner power or economic rewards. No group has a God-given or hereditary interest in any field, and any effort to claim infallibility should be rigorously resisted.

On the current scene, signs of academic imperialism may be discerned in the attempts of the older university departments to maintain power and to control the curriculum in order to claim the largest numbers of students. The rise of the social sciences has been handicapped by the unwillingness of entrenched departments to surrender large blocks of financial resources and time to the younger departments. And within one department, there may be a struggle between different fields for priority as has occurred and been described publicly in several American departments of psychology. A current focus of conflict is between basic and applied science for power, with basic science claiming that it has scientific priority and applied science claiming that it must satisfy the pressure of social demands.

In the clinical field, the largest scale potential conflict exists between organized medicine and other contenders for control of the healing arts. Supported by powerful resources, the medical profession has thus far been very successful in controlling clinical practice and relegating other professions to the position of being ancillary technicians. The struggle promises to be particularly brisk in the field of psychotherapy where psychiatry is only one competitor against psychologists, educators, clergymen and social workers each of whom may claim competency in certain areas. In areas where the conflict between medical and lay psychotherapists has been strongest, it is frequently possible to discern economic motives being an important issue at stake. In private conversations, some medical specialists decry the loss of income which would occur if large numbers of lay therapists suddenly appeared. Fortunately, this area of conflict is being peacably resolved by inter-disciplinary negotiating committees.

An interesting example of professional imperialism lies in the efforts of the psychiatric profession to assume control of the mental hygiene movement as their particular responsibility. We hear prominent psychiatrists lecturing to educators and others on the topic of what psychiatry has to offer in the mental hygiene field, as if assuming that psychiatry has prior responsibility and was the only discipline having something to offer in the preventative field. The educators tend to listen to such influence with tongue in cheek and promptly go out and establish their own child care agencies oriented in an educational rather than a psychiatric direction.

Imperialism becomes more complex when several disciplines band together to assert domination over a field. For example, the Trinity of psychiatrist, psychologist and social worker has been constituted as a team to take over clinical work. Such a

pattern of professional organization tends to become frozen, and to exclude other groups with potential contributions to make.

On even more insidious levels is the attempt of specialists within a specialty to construct an empire for themselves. This has occurred in the psychological field in two patterns. In many areas, psychoanalysts have gained control over local facilities to the exclusion of non-analytically trained specialists, and the only students who can gain a foothold are those who enroll for analytic training. A second type of potential imperialism occurs when a small group of specialists within a specialty designate themselves as sort of an "elite" guard, and proceed to infiltrate the official organization of the profession in order to control power and policies. Fortunately, such movements stimulate the organization of opposing groups which militantly fight for democratic control and operation of professional activities.

Imperialism has no place in scientific professions because it implies the authoritarianism and dogma which are the greatest blocks to scientific progress. Scientific fields must be kept open to all who possess equal competence and have any contribution to offer. And certainly, scientific knowledge is evolving so rapidly that it is dangerous for adherents of any particular theory or system of practice to have their methods designated as the only approved techniques in the field. The attempt to protect the *status quo* by freezing current patterns and practices is a dangerous step along the road of professional imperialism.

F. C. T.

THE HOWARD INK BLOT TEST

A NEW PROJECTIVE TECHNIQUE

- **Derived experimentally by James W. Howard, Ph.D.**
- **An individual test consisting of 12 plates 8½" x 11". The blots are somewhat larger than the Rorschach and with a wider range of stimulating features. Blots of varied colors selected from an experimental sample of 10,000 blots.**
- **Not intended to be a parallel test to the Rorschach. Statistical findings from a "normal" group show the test to differ markedly from the Rorschach in ways which should indicate greater sensitivity.**
- **Selection and arrangement of blots made on the basis of evidence of character, degree and spread of stimulation.**
- **Data from normative sample indicate more frequent light determined responses, higher incidence of color responses with different C, CF and FC relationships, more movement responses and fewer animal responses. Greater diagnostic sensitivity with wider range of response.**

Price: $12.50 (In U. S. and Canada; foreign postage $.50 extra)

HOWARD INK BLOT TEST MANUAL

MONOGRAPH SUPPLEMENT NO. 10 FROM JULY 1953 ISSUE

Price $2.00

JOURNAL OF CLINICAL PSYCHOLOGY
5 Pearl Street Brandon, Vermont

BOOK NOTICES

SHELDON, W. H. *Atlas of Men.* New York: Harper, 1954, pp. 357. $10.00

This atlas contains photographs representing front, side and rear views of 1,175 men selected as examples of the 88 known somatotypes as classified by Sheldon's system. Tabular materials present the approximate incidence of each somatotype; their arrangement according to rarity; standard average weights for height and age; mean statures for the somatotypes; means, standard deviations and percentiles for the primary components in man; and distributions of somatotypes on height over cube root of weight for various age groups. These data make it possible to objectively determine any male's somatotype as operationally defined by Sheldon's system. The photographs illustrate the various somatotypes clearly, and will be of great assistance to students of somatotypy. Further assistance is provided by a nomograph for determining height over the cube root of weight when height is known in inches and weight in pounds and resulting in a ratio which, when related to stables of the distributions of the somatotypes for various age groups, facilitates the approximation of the particular type involved. Our only criticism of this atlas refers to the highly speculative and impressionistic textual descriptions of the various somatotypes which go beyond anything which is scientifically known and represent only the interpretations of the author.

JONES, MARSHALL R. (Ed.) *Nebraska Symposium on Motivation: 1955.* Lincoln, Neb.: University of Nebraska Press, 1955, p. 274. $3.00

University of Nebraska Press, 1955, pp274. $3.00. A series of essays on motivation by Abraham Maslow, David C. McClelland, James Olds, Helen Peak, Julian B. Rotter and Paul Thomas Young.

ABRAMSON, H. A. (Ed.) *Problems of Consciousness.* New York: Josiah Macy, Jr.

Foundation, 1955, pp. 180. $3.50. Proceedings of the fifth conference sponsored by the Josiah Macy, Jr. Foundation and held on March 22-24, 1954 on problems of consciousness.

SCHNEIDERS, A. A. *Personal Adjustment and Mental Health.* New York: Rinehart, 1955, pp. 587. $5.00

This textbook succeeds far better than most in the field in integrating basic science psychology with abnormal psychology and psychopathology in an approach to mental hygiene and personal adjustment. A preliminary section of two chapters outlines approaches to the study of adjustment. Part II includes seven chapters on the nature and conditions of adjustment. Part III outlines the principal patterns and mechanisms of adjustment. Part IV presents two chapters describing varieties of adjustment in personal, social, vocational and marital areas, and a concluding section of two chapters discusses mental hygiene and treatment. This book is not as heavily weighted with psychiatric material as previous textbooks on mental hygiene.

LEMKAU, P. V. *Mental Hygiene in Public Health.* New York: McGraw-Hill, 1955, pp. 486. $8.00. Second edition.

BORNEMISZA, S. T. *The Unified System Concept of Nature.* New York: Vantage, 1955, p. 137. $3.00

BRENNER, M. D. *An Elementary Textbook of Psychoanalysis.* New York: International Universities Press, 1955, pp. 219. $4.00

WEINSTEIN, E. A. *Denial of Illness.* Springfield, Ill.: C. C. Thomas, 1955, pp. 166. $4.75

COLE, W. G. *Sex in Christianity and Psychoanalysis.* New York: Oxford University Press, 1955, pp. 329. $4.00

BOISEN, A. T. *Religion in Crisis and Custom.* New York: Harper, 1955, pp. 271. $4.00

DARWIN, C. *Expression of the Emotions in Man and Animals.* New York: Philosophical Library, 1955, pp. 372. $6.00

HARROWER, M. (Ed.) *Medical and Psychological Teamwork in the Care of the Chronically Ill.* Springfield, Ill.: C. C. Thomas, 1955, pp. 232. $5.75

ROHEIM, G. *Magic and Schizophrenia.* New York: International Universities Press, 1955, pp. 230. $4.50

SELL, D. E. (Ed.) *Manual of Applied Correctional Psychology.* Columbus, Ohio: Department of Mental Hygiene and Correction, 1955, pp. 68.

ROBACK, A. A. (Ed.) *Present-day Psychology.* New York: Philosophical Library, 1955, pp. 995. $12.00

LINN, L. *A Handbook of Hospital Psychiatry.* New York: International Universities Press, 1955, pp. 560. $10.00

THE SOUTHARD SCHOOL

Intensive individual psychotherapy in a residential school, for children of elementary school age with emotional and behavior problems.

THE MENNINGER CHILDREN'S CLINIC

Outpatient psychiatric and neurologic evaluation and consultation for infants and children to eighteen years.

Department of Child Psychiatry

THE MENNINGER FOUNDATION

J. Cotter Hirschberg, M. D., Director **Topeka, Kansas; Telephone 3-6494**

A STIMULATING REASSESSMENT OF

Community Programs for Mental Health

Editors
RUTH KOTINSKY
HELEN L. WITMER

Contributors
BARBARA BIBER
H. E. CHAMBERLAIN
SOL W. GINSBURG
ROBERT R. HOLT
LOUISA P. HOWE
MARIE JAHODA
ELIZABETH DESCHWEINITZ
EDITH MILLER TUFTS

THEORY, PRACTICE, AND EVALUATION

This book provides an overall view of mental health promotion—its underlying theory, typical practice, and problems of evaluation. Each of the papers focuses from its own vantage point on current efforts to maintain the mental health of all at an optimum level on a community-wide basis.

The authors of these papers discovered interesting programs of mental health education and pioneer attempts to translate mental health principles into professional practice. They also found, however, that the theoretical basis of the various endeavors in this area is inadequately developed, that evaluation is a thing of the future, and that the very definition of mental health promotion poses serious difficulties. Their findings and analyses are therefore of the greatest importance, for they will inevitably stimulate further thought and action in this fertile area of public health and welfare.

Teachers, students, and practitioners of psychology and psychiatry, and of social work and nursing, and research workers in the social sciences generally, will all value this book. 384 pages. $5.00

ANOTHER IMPORTANT STUDY

PSYCHOMOTOR ASPECTS OF MENTAL DISEASE

An Experimental Study. By H. E. King. 1954. 199 pages. 21 figures. 21 tables. $3.50

COMMONWEALTH FUND BOOKS
Through Your Bookstore, or from
HARVARD UNIVERSITY PRESS
Cambridge 38 Massachusetts

d is for distance

D is for Devereux Schools

which know that the shortest distance between two points is not always a straight line -- but the route which affords the fewest "intervening opportunities." (Stouffer: Am. Soc. Rev. 5:845, 1940)

The business of Devereux Schools is to structure "intervening opportunities," in accordance with each student's needs individually -- needs for proximate motivation and rewards and needs for remote-goal attainment. In more than a score of geographically separated schools, such structuring is worked out objectively for either the emotionally confused or the impaired in terms of CNS function.

Detailed information is available to the profession, without obligation, in response to a letter or post card addressed to

JOHN M. BARCLAY
Director of Development
The Devereux Foundation
Devon, Pennsylvania

Devereux Schools

UNDER THE DEVEREUX FOUNDATION

Helena T. Devereux, Director

Santa Barbara, California **Devon, Pennsylvania**